

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:27:19 ; Search time 21.78 Seconds

(without alignments)  
405.329 Million cell updates/sec

Title: US-09-853-526-70

Perfect score: 228  
Sequence: 1 MKYLLPSVVLGTAFTYVLA.....NYLDAIVDTVVEKGDG 228

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size : 0

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	228	100.0	353	1 PICE_HUMAN	Q9nqg2 homo sapien
2	32	14.0	354	1 PICE_MOUSE	Q9d1e8 mus musculu
3	9	3.9	237	1 PISC_HELPY	Q9zjn8 heliocobacte
4	9	3.9	240	1 PISC_HELPY	Q25903 heliocobacte
5	8	3.5	396	1 YD18_YEAST	Q12185 saccharomyc
6	8	3.5	611	1 HSCA_BUGAI	P57660 buchnera ap
7	7	3.1	57	1 YJDO_ECOLI	P58038 escherichia
8	7	3.1	194	1 PURI_LACCA	P35853 lactobacilli
9	7	3.1	208	1 HBGF_CERAE	Q09118 cercopithec
10	7	3.1	208	1 HBGF_HUMAN	Q09075 homo sapien
11	7	3.1	208	1 HBGF_PIG	Q01580 sus scrofa
12	7	3.1	250	1 PLSC_BORBU	Q06175 rattus norv
13	7	3.1	250	1 PLSC_BORBU	Q05988 borrellia bu
14	7	3.1	250	1 PLSC_BORBU	Q23449 caenorhabdi
15	7	3.1	286	1 PSCD_CAEEL	P32263 saccharomyc
16	7	3.1	308	1 PISC_COCNU	Q42670 coccos nucif
17	7	3.1	445	1 XYLA_BACSU	P13860 phenocochae
18	7	3.1	516	1 GUX1_PHACH	P10480 nocardia la
19	7	3.1	520	1 CMCH_NOCUA	P05080 nocardia la
20	7	3.1	532	1 GSI_NEUCR	P38678 neurospora
21	7	3.1	677	1 Y593_TREPA	Q08362 treponema p
22	7	3.1	860	1 CH12_COCIM	P54197 coccioidoide
23	7	3.1	904	1 Y002_CAEEL	Q09228 caenorhabdi
24	7	3.1	2485	1 PTMD_HUMAN	Q12923 homo sapien
25	6	2.6	25	1 PCW1_PACGO	P82423 pachycondyl
26	6	2.6	25	1 PCW2_PACGO	P82424 pachycondyl
27	6	2.6	76	1 BBL1_SCHCO	P78742 schizophy11
28	6	2.6	83	1 RL23_HALGA	Q06842 halobacteri
29	6	2.6	87	1 YIMC_BPPH1	P10435 bacteriophi
30	6	2.6	88	1 PMRD_ECOLI	P37590 escherichia
31	6	2.6	99	1 Y4OL_RHISN	P55597 rhizobium s
32	6	2.6	99	1 PROL_METTA	Q05832 methanococc
33	6	2.6	101	1 THST_THERS	Q9nbw4 thiermyzon

## ALIGNMENTS

RESULT	1	STANDARD:	PRT:	353 AA.
PLCE_HUMAN				
ID	PLCE_HUMAN			
AC	Q9NIO2; Q9BQ64;			
DT	01-MAR-2002 (Rel. 41, Created)			
DT	01-MAR-2002 (Rel. 41, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	1-acyl-sn-glycerol-3-phosphate acyltransferase epsilon (EC 2.3.1.51)			
DE	(1-AGP acyltransferase 5) (1-AGPAT 5) (lysophosphatidic acid			
DE	acyltransferase-epsilon) (LPAT-epsilon) (1-acylglycerol-3-phosphate			
DE	O-acyltransferase 5).			
GN	AGPAT5.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OX	Mammalia; Euteria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_Taxid-9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Leung D.W.;			
RT	"Cloning and expression of LPAT-epsilon."			
RL	Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.			
RP	SEQUENCE FROM N.A.			
RA	Cohen D., Chumakov I., Blumenfeld M., Bougueleret L.;			
RL	Patent number WO9932644, 01-JUL-1999.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE-21154917; PubMed-11230166;			
RC	Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glasel S.,			
RA	Ansorge W., Boecker M., Bloeker H., Bauersachs S., Blum H.,			
RA	Lauber J., Duesterhoeft A., Beyer A., Koehler K., Strack N.,			
RA	Mewes H.-W., Ottenwelder B., Obermaier B., Tampe J., Heubner D.,			
RT	Wambut R., Korn B., Klein M., Poustka A.;			
RT	"Towards a catalog of human genes and proteins: sequencing and			
RT	analysis of 500 novel complete protein coding human cDNAs."			
RL	Genome Res. 11:422-435(2001).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RA	Tissue-Placenta;			
RC	Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,			
RA	Nishikawa T., Nagai K., Sugano S., Shitatori A., Sudo H.,			
RA	Yagashima M., Hosoliri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,			
RA	Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,			
RA	Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,			
RA	Yamamoto J., Wakematsu A., Nakamura Y., Nishihara K., Masuno Y.,			
RA	Ninomiya K., Iwyanagi T.;			
RT	"NEO human cDNA sequencing project."			
RT	Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.			
CC	-1- FUNCTION: CONVERTS LYSPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC			
CC	ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION (BY			
CC	STIMULATORY).			
CC	-1- CATALYTIC ACTIVITY: AYL-CoA + 1-acyl-sn-glycerol 3-phosphate =			
CC	CoA + 1,2-diacyl-sn-glycerol 3-phosphate.			
CC	-1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).			

CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
CC ACYLTRANSFERASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC at the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: AF375789; AAK54809.1; ALT\_INT.  
CC EMBL: AL136587; CAB66222.1; ALT\_INT.  
CC EMBL: AK002072; BAA32069.1; -  
CC InterPro: IPR002123; Acyltransferase.  
CC Pfam: PF01553; Acyltransferase: 1.  
CC Phospholipid biosynthesis; Transferase; Acyltransferase;  
CC Transmembrane.  
CC TRANSMEM 7 29 POTENTIAL.  
CC TRANSMEM 44 66 POTENTIAL.  
CC TRANSMEM 334 351 POTENTIAL.  
CC CONFLICT 145 145 L -> V (IN REF. 2).  
CC SQUENCE 353 AA; 40813 MW; A05B1FA246CE1B64 CRC64;

Query Match 100.0%; Score 228; DB 1; Length 353;  
Best Local Similarity 100.0%; Pred. No. 1.4e-227;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRYLLPSVLLGTAPYVYLANGVRLSAPLPARFYQALDDRLCYQSNVLEFFENYTG 60  
DB 1 MRYLLPSVLLGTAPYVYLANGVRLSAPLPARFYQALDDRLCYQSNVLEFFENYTG 60  
OY 61 VOILLYDDLPKKNKNIYILANHOSTVMDIADIAIRQNALGHRYVYKGLKMLFYYGC 120  
DB 61 VOILLYDDLPKKNKNIYILANHOSTVMDIADIAIRQNALGHRYVYKGLKMLFYYGC 120  
OY 121 YFAOHGIGYVRSKAFNEKEKRNKLSQSYVDAGTPMVLVIEPEGTRVNEQTKVLSAQAF 180  
DB 121 YFAOHGIGYVRSKAFNEKEKRNKLSQSYVDAGTPMVLVIEPEGTRVNEQTKVLSAQAF 180  
OY 181 AAOGLAVLKHVLTPIKATIHAFDCKMKNYLDATYDVTYVYEGKDDG 228  
DB 181 AAOGLAVLKHVLTPIKATIHAFDCKMKNYLDATYDVTYVYEGKDDG 228

RESULT 2  
PLSC\_MOUSE STANDARD; PRT; 354 AA.  
AC Q9DIE8;  
DT 01-MAR-2002 (Rel. 41, Created)  
DT 01-MAR-2002 (Rel. 41, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE 1-acyl-sn-glycerol-3-phosphate acyltransferase epsilon (EC 2.3.1.51)  
DE (1-AGP acyltransferase 5) (1-AGPAT 5) (lysophosphatidic acid  
DE acyltransferase-epsilon) (1-PAT-epsilon) (1-acylglycerol-3-phosphate  
DE O-acyltransferase 5).  
GN AGPAT5.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N. A.  
RC STRAIN=C57BL/6J; TISSUE=Embryo;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Aikawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
RA Aikawa K., Iwano M., Nishi K., Kiyosawa H., Kondo S., Yamana K. I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H. A., Ashburner M., Batilov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Knehl P., Lewis S., Matsuo Y., Mikaido I., Pesole G., Quackenbush J.,

RA Schirml L. M., Staabli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M. F.,  
RA Brownstein M. J., Butt C., Fletcher C., Fujita M., Gariboldi M.,  
RA Cusumano S., Hill D., Hofmann M., Hume D. A., Kamiya M., Lee N. H.,  
RA Lyons F., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Norioka P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K. F.,  
RA Suzuki H., Toyooka K., Wang K. H., Weltz C., Whitaker C., Wilmink L.,  
RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,  
RA Hashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection";  
RL Nature 403:680-690(2001).  
CC -1- FUNCTION: CONVERTS LYSOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC  
CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION (BY  
CC SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate -  
CC CoA + 1,2-diacyl-sn-glycerol 3-phosphate.  
CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
CC ACYLTRANSFERASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.  
CC at the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: AK003649; BAB22915.1; -  
CC MGI: MGI:1915880; Agpat5.  
CC InterPro: IPR002123; Acyltransferase.  
CC Pfam: PF01553; Acyltransferase: 1.  
CC Phospholipid biosynthesis; Transferase; Acyltransferase;  
CC Transmembrane.  
CC TRANSMEM 7 29 POTENTIAL.  
CC TRANSMEM 44 66 POTENTIAL.  
CC TRANSMEM 335 352 POTENTIAL.  
CC SQUENCE 354 AA; 40943 MW; 447EEF924B91E800 CRC64;

Query Match 14.0%; Score 32; DB 1; Length 354;  
Best Local Similarity 100.0%; Pred. No. 3e-25;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 174 LSASQAFAPAGRLAVLKHVLTPIKATIHAFD 205  
DB 174 LSASQAFAPAGRLAVLKHVLTPIKATIHAFD 205

RESULT 3  
PLSC\_HELPJ STANDARD; PRT; 237 AA.  
AC Q9ZJN8;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE 1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51) (1-AGP  
DE acyltransferase) (1-AGPAT) (lysophosphatidic acid acyltransferase)  
DE (LPAAT).  
GN PLSC OR JHP1267.  
OS Helicobacter pylori 399 (Campylobacter pylori 399).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=85963;  
RN [1]  
RP SEQUENCE FROM N. A.  
RX MEDLINE=99120557; PubMed=9923682;  
RA Alm R. A., Ling L. S. L., Moll D. T., King B. L., Brown E. D., Doig P. C.,  
RA Smith D. R., Noonan B., Guild B. C., deJonge B. L., Carmel G.,  
RA Tummlino P. J., Caruso A., Uria-Nickelsen M., Mills D. M., Ives C.,

RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Voyis G.F.,  
 RA Trust T.J.;  
 RT "Genomic sequence comparison of two unrelated isolates of the human  
 RT gastric pathogen *Helicobacter pylori*.";  
 RL Nature 397:176-180(1999).  
 CC -1- FUNCTION: CONVERTS LYSPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC  
 CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION.  
 CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate =  
 CC CoA + 1,2-diacyl-sn-glycerol 3-phosphate.  
 CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.  
 CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
 CC ACYLTRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AE001550; AAD06852.1; -;  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KW Phospholipid biosynthesis; Transferase; Acyltransferase;  
 KW Inner membrane; Complete proteome  
 SQ SEQUENCE 237 AA; 2718 MM; E10F517D42A1731F CRC64;

Query Match 3.9%; Score 9; DB 1; Length 237;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 LVIFPEGTR 165  
 |||||  
 Db 142 LVIFPEGTR 150

RESULT 4  
 PLSC\_HELPY STANDARD: PRT; 240 AA.  
 AC 025903;  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE 1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51) (1-AGP  
 DE acyltransferase) (1-AGPAT) (Lysophosphatidic acid acyltransferase)  
 DE (LPAAT).  
 GN PLSC OR HP1348.  
 OS *Helicobacter pylori* (Campylobacter *pylori*).  
 OC Bacteria; Proteobacteria; epsilon subdivision; *Helicobacter* group;  
 OC *Helicobacter*.  
 OX NCBI\_TaxID=210;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=26695 / ATCC 700392;  
 RX MEDLINE=97394467; PubMed=9252185;  
 RA Toml J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,  
 RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,  
 RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,  
 RA Loftus B., Fitzgerald D., Dodson R., Khaila S.G., Glodek A.,  
 RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,  
 RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,  
 RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,  
 RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,  
 RA Venter J.C.;  
 RT "The complete genome sequence of the gastric pathogen *Helicobacter*  
 RT *pylori*.";  
 RL Nature 388:539-547(1997).  
 CC -1- FUNCTION: CONVERTS LYSPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC  
 CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION.  
 CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate =

CC CoA + 1,2-diacyl-sn-glycerol 3-phosphate.  
 CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.  
 CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
 CC ACYLTRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AE000636; AAD08393.1; -;  
 DR TIGR: HP1348; -;  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KW Phospholipid biosynthesis; Transferase; Acyltransferase;  
 KW Inner membrane; Complete proteome  
 SQ SEQUENCE 240 AA; 27745 MM; 22B05D0EB190BBD CRC64;

Query Match 3.9%; Score 9; DB 1; Length 240;  
 Best Local Similarity 100.0%; Pred. No. 0.12;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 LVIFPEGTR 165  
 |||||  
 Db 142 LVIFPEGTR 150

RESULT 5  
 YD18\_YEAST STANDARD: PRT; 396 AA.  
 AC 012185;  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DE Hypothetical 45.9 kDa protein in KCS1-GCV1 intergenic region.  
 GN YDR018C OR YD9335.04C OR P2F396.  
 OS *Saccharomyces cerevisiae* (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; *Saccharomycotina*; *Saccharomycetes*;  
 OC *Saccharomycetales*; *Saccharomycetaceae*; *Saccharomyces*.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C / AB972;  
 RA Dedman K., Brown D., Hamlyn N., Bowman S., Barrell B.G.,  
 RL Submitted (MAY-1995) to the EMBL/Genbank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97051598; PubMed=8896275;  
 RA Elde L.G., Sander C., Prydz H.;  
 RT "Sequencing and analysis of a 35.4 kb region on the right arm of  
 RT chromosome IV from *Saccharomyces cerevisiae* reveal 23 open reading  
 RT frames.";  
 RL Yeast 12:1085-1090(1996).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
 CC ACYLTRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: 274314; CAA98838.1; -;  
 DR EMBL: X95966; CAA65210.1; -;

DR EMBL: 249770; CAAB9643.1; -.  
 DR SGD: S0002425; YDR018C.  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KM Hypothetical protein: Phospholipid biosynthesis; Transferase;  
 KM Acyltransferase; Transmembrane.  
 FT TRANSMEM 27 47 POTENTIAL.  
 FT TRANSMEM 69 69 POTENTIAL.  
 FT TRANSMEM 123 143 POTENTIAL.  
 FT TRANSMEM 372 392 POTENTIAL.  
 SO SEQUENCE 396 AA; 45938 MW; 9F55AB56C8FD44DD CRC64;

Query Match 3.5%; Score 8; DB 1; Length 396;  
 Best Local Similarity 100.0%; Pred. No. 1.9;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 211 LDAIYDVT 218  
 DB 265 LDAIYDVT 272

RESULT 6  
 HSCA\_BUCAI STANDARD; PRT; 611 AA.  
 ID HSCA\_BUCAI  
 AC P57660;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Chaperone protein hscA homolog.  
 GN HSCA OR BU605.  
 OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum  
 symbiotic bacterium).  
 OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.  
 OX NCBI\_TaxID=118099;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TOYO 1998;  
 RX MEDLINE=20445173; PubMed=10993077;  
 RA Shigenobu S., Matanabe H., Hattori M., Sakaki Y., Ishikawa H.;  
 RT "Genome sequence of the endocellular bacterial symbiont of aphids  
 Buchnera sp. AS.";  
 RL Nature 407:81-86(2000).  
 CC -1- FUNCTION: PROBABLE CHAPERONE. HAS A LOW INTRINSIC ATPASE ACTIVITY  
 WHICH IS MARKEDLY STIMULATED BY HSCB (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: AP001119; BAB13289.1; -.  
 DR InterPro: IPR001023; HSP70.  
 DR Pfam: PF00012; HSP70; 1.  
 DR PRINTS: PR00301; HEATSHOCK70.  
 DR PROSITE: PS00287; HSP70\_1; 1.  
 DR PROSITE: PS00329; HSP70\_2; 1.  
 DR PROSITE: PS01036; HSP70\_3; 1.  
 KM Chaperone; ATP-binding; Complete proteome.  
 SO SEQUENCE 611 AA; 69837 MW; 32B567C53073082A CRC64;

Query Match 3.5%; Score 8; DB 1; Length 611;  
 Best Local Similarity 100.0%; Pred. No. 2.8;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RYLLPSVY 9  
 DB 42 RYLLPSVY 49

RESULT 7  
 YJDO\_ECOLI STANDARD; PRT; 57 AA.  
 ID YJDO\_ECOLI  
 AC P58038;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Hypothetical protein yjdo.  
 GN YJDO OR B4128.1 OR 25731 OR EC55110.1.  
 OS Escherichia coli, and  
 OS Escherichia coli O157:H7.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562, 83334;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12.";  
 RL Science 277:1453-1474(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;  
 RX MEDLINE=21074935; PubMed=11206551;  
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,  
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
 RA Welch R.A., Blattner F.R.;  
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";  
 RL Nature 409:529-533(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / RIMD 0509952;  
 RX MEDLINE=21156231; PubMed=11258796;  
 RA Hayashi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K.,  
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,  
 RA Kihara S., Shiba T., Hattori M., Shingawa H.;  
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli  
 O157:H7 and genomic comparison with a laboratory strain K-12.";  
 RL DNA Res. 8:11-22(2001).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: STRONG, TO E.COLI YDCX.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: AE000485; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL: AE005646; AAG59328.1; -.  
 DR EMBL: AP002568; -; NOT\_ANNOTATED\_CDS.  
 DR EcoGene: EG14342; yjdo.  
 KM Hypothetical protein: Transmembrane; Complete proteome.  
 FT TRANSMEM 7 27 POTENTIAL.  
 FT TRANSMEM 37 57 POTENTIAL.  
 SO SEQUENCE 57 AA; 6555 MW; A3670A19500F75D6 CRC64;

Query Match 3.1%; Score 7; DB 1; Length 57;  
 Best Local Similarity 100.0%; Pred. No. 4;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



QY 25 RLISAFL 31  
 |||||  
 DB 33 RLISAFL 39

RESULT 8  
 ID PURI\_LACCA STANDARD; PRT; 194 AA.  
 AC P35853;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Amidophosphoribosyltransferase precursor (EC 2.4.2.14) (Glutamine  
 phosphoribosylpyrophosphate amidotransferase) (ATPase) (GPRATase)  
 DE (Fragment).  
 GN PURP.  
 OS Lactobacillus casei.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;  
 OC Lactobacillus.  
 OX NCBI\_TaxID=1582;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93012962; PubMed=1398079;  
 RA Gu Z.-M., Martindale D.W., Lee B.H.;  
 RT "Isolation and complete sequence of the purL gene encoding PCAM  
 synthase II in Lactobacillus casei.";  
 RL Gene 119:123-126(1992).  
 RN [2]  
 RP ERRATUM.  
 RX MEDLINE=94040790; PubMed=8224889;  
 RA Gu Z.-M., Martindale D.W., Lee B.H.;  
 RL Gene 133:147-147(1993).  
 CC -1- CATALYTIC ACTIVITY: 5-phospho-beta-D-ribosylamine + diphosphate +  
 L-glutamate -> L-glutamine + 5-phospho-alpha-D-ribose 1-diphosphate +  
 H(2O).  
 CC -1- PATHWAY: FIRST STEP IN DE NOVO PURINE BIOSYNTHESIS.  
 CC -1- SIMILARITY: THE GATASE DOMAIN BELONGS TO TYPE-2 GLUTAMINE  
 AMIDOTRANSFERASES.  
 CC -1- SIMILARITY: IN THE C-TERMINAL SECTION, BELONGS TO THE  
 PURINE/PRIMIDINE PHOSPHORIBOSYLTRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M85265; AAC36948.1; .  
 DR PIR: PC1136; PC1136.  
 DR HSSP: P00497; 1A00.  
 DR MEROPS: C44.001; .  
 DR InterPro: IPR000583; GATase\_2.  
 DR InterPro: IPR002375; Pur\_pyr\_pr\_transf.  
 DR Pfam: PF00310; GATase\_2; 1.  
 DR PROSITE: PS00103; PUR\_PYR\_PR\_TRANSFER; PARTIAL.  
 DR PROSITE: PS00443; GATASE\_TYPE\_II; 1.  
 KW Purine biosynthesis; Transferase; Glycosyltransferase.  
 FT PROPEP 1 11 BY SIMILARITY.  
 FT CHAIN 12 >194 AMIDOPHOSPHORIBOSYLTRANSFERASE.  
 FT ACU\_SITE 12 12 GATASE (INVOLVED IN SUBSTRATE BINDING AND  
 CATALYSIS) (BY SIMILARITY).  
 FT NON\_TER 194 194  
 SO SEQUENCE 194 AA; 21144 MW; 4A78CB5365D5EC CRC64;

Query Match 3.1%; Score 7; DB 1; Length 194;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 100 ALGHVRY 106

DB 79 ALGHVRY 85  
 |||||

RESULT 9  
 ID HBGF\_CERAE STANDARD; PRT; 208 AA.  
 AC Q09118;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Heparin-binding EGF-like growth factor precursor (HB-EGF) (HBEGF)  
 DE (Diphtheria toxin receptor) (DT-R).  
 GN DTR OR HBGF.  
 OS Cercopithecus aethiops (Green monkey) (Griwet).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;  
 OC Cercopithecinae; Cercopithecus.  
 OX NCBI\_TaxID=9534;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92298386; PubMed=1606612;  
 RA Naglich J.G., Metherall J.E., Russel D.W., Bideals L.;  
 RT "Expression cloning of a diphtheria toxin receptor: Identity with a  
 heparin-binding EGF-like growth factor precursor.";  
 RL Cell 69:1051-1061(1992).  
 RN [2]  
 RP TOXIN-BINDING DOMAIN.  
 RX MEDLINE=95126975; PubMed=7826391;  
 RA Hooper K.P., Bideals L.;  
 RT "Localization of a critical diphtheria toxin-binding domain to the C-  
 terminus of the mature heparin-binding EGF-like growth factor region  
 of the diphtheria toxin receptor.";  
 RL Biochem. Biophys. Res. Commun. 206:710-717(1995).  
 CC -1- FUNCTION: MAY BE INVOLVED IN MACROPHAGE-MEDIATED CELLULAR  
 PROLIFERATION. IT IS MITOGENIC FOR FIBROBLASTS AND SMOOTH MUSCLE  
 BUT NOT ENDOTHELIAL CELLS. IT IS ABLE TO BIND EGF RECEPTORS WITH  
 HIGHER AFFINITY THAN EGF ITSELF AND IS A FAR MORE POTENT MITOGEN  
 FOR SMOOTH MUSCLE CELLS THAN EGF (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MATURE HB-EGF IS  
 RELEASED INTO THE EXTRACELLULAR SPACE AND PROBABLY BINDS TO A  
 RECEPTOR (BY SIMILARITY).  
 CC -1- PTM: O-GLYCOSYLATED (BY SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M93012; .; NOT\_ANNOTATED\_CDs.  
 DR HSSP: Q99075; 1XDY.  
 DR InterPro: IPR000561; EGF-like.  
 DR Pfam: PF00008; EGF; 1.  
 DR SMART: SM00181; EGF; 1.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 KW Growth factor; Heparin-binding; signal; EGF-like domain;  
 KW Transmembrane; Glycoprotein; Receptor.  
 FT SIGNAL 1 19  
 FT PROPEP 20 62  
 FT CHAIN 63 148  
 FT PROPEP 149 208  
 FT DOMAIN 20 160  
 FT TRANSMEM 161 184  
 FT DOMAIN 185 208  
 FT DOMAIN 104 144  
 FT CARBOHYD 75 75  
 FT CARBOHYD 85 85  
 FT DISULFID 108 121  
 FT -----  
 FT BY SIMILARITY.  
 FT HEPARIN-BINDING EGF-LIKE GROWTH FACTOR.  
 FT C-TERMINAL (POTENTIAL).  
 FT EXTRACELLULAR (POTENTIAL).  
 FT POTENTIAL.  
 FT CYTOPLASMIC (POTENTIAL).  
 FT EGF-LIKE.  
 FT O-LINKED (GALNAC. . .) (BY SIMILARITY).  
 FT O-LINKED (GALNAC. . .) (BY SIMILARITY).  
 FT BY SIMILARITY.

```

FT DISULFID 116 132 BY SIMILARITY.
FT DISULFID 134 143 BY SIMILARITY.
SQ SEQUENCE 208 AA; 22985 MW; 8D108289AD0485AE9 CRC64;

Query Match 3.1%; Score 7; DB 1; Length 208;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 LPSVVL 10
Db 3 LPSVVL 9

RESULT 10
HBGF_HUMAN STANDARD; PRT; 208 AA.
AC 099075;
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Heparin-binding EGF-like growth factor precursor (HB-EGF) (HBEGF)
DE (Diphtheria toxin receptor) (DTR-R).
GN DTR OR HBGF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 73-93.
RC TISSUE=Macrophage;
RX MEDLINE=91157008; PubMed=1840698;
RT Higashiyama S., Abraham J.A., Miller J., Fiddes J.C., Klagsbrun M.;
RT "A heparin-binding growth factor secreted by macrophage-like cells
RT that is related to EGF.";
RL Science 251:936-939(1991).
RN [2]
RP SEQUENCE FROM N.A.
RA Kimerly W., Bondoc M., Cheng J., Connolly K.S., Gunning K.M.,
RA Kader K., Miguel T., Miller C., Pittluck S., Pollard M., Rojeski H.,
RA Subramanian S., Martin C.H.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 63-141 AND 143-148, AND CARBOHYDRATE-LINKAGE SITES.
RC TISSUE=Histocytic lymphoma;
RX MEDLINE=92210596; PubMed=1556128;
RA Higashiyama S., Lau K., Besner G.E., Abraham J.A., Klagsbrun M.;
RT "Structure of heparin-binding EGF-like growth factor. Multiple forms,
RT primary structure, and glycosylation of the mature protein.";
RL J. Biol. Chem. 267:6205-6212(1992).
RN [4]
RP TOXIN-BINDING DOMAIN.
RX MEDLINE=95138082; PubMed=7836353;
RA Mitamura T., Higashiyama S., Taniguchi N., Klagsbrun M., Mekada E.;
RT "Diphtheria toxin binds to the epidermal growth factor (EGF)-like
RT domain of human heparin-binding EGF-like growth factor/diphtheria
RT toxin receptor and inhibits specifically its mitogenic activity.";
RL J. Biol. Chem. 270:1015-1019(1995).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.65 ANGSTROMS) OF 73-147 IN COMPLEX WITH TOX.
RX MEDLINE=98324089; PubMed=9559904;
RA Louie G.V., Yang W., Bowman M.E., Choe S.;
RT "Crystal structure of the complex of diphtheria toxin with an
RT extracellular fragment of its receptor.";
RL Mol. Cell 1:67-78(1997).
RN [6]
RP -1- FUNCTION: MAY BE INVOLVED IN MACROPHAGE-MEDIATED CELLULAR
RP PROLIFERATION. IT IS MITOGENIC FOR FIBROBLASTS AND SMOOTH MUSCLE
RP BUT NOT ENDOTHELIAL CELLS. IT IS ABLE TO BIND EGF RECEPTORS WITH
RP HIGHER AFFINITY THAN EGF ITSELF AND IS A FAR MORE POTENT MITOGEN
RP FOR SMOOTH MUSCLE CELLS THAN EGF.
RP -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MATURE HB-EGF IS
RP RELEASED INTO THE EXTRACELLULAR SPACE AND PROBABLY BINDS TO A
RP RECEPTOR.

```

```

CC -1- PTM: SEVERAL N-TERMINI HAVE BEEN IDENTIFIED BY DIRECT SEQUENCING.
CC THE FORMS WITH N-TERMINI 63, 73 AND 74 HAVE BEEN TESTED AND FOUND
CC TO BE BIOLOGICALLY ACTIVE.
CC -1- PTM: O-GLYCAN ATTACHMENT SITES WERE DETERMINED BY EDMAN
CC DEGRADATION, O-GLYCANASE DIGEST SUGGESTS MUCIN-TYPE GLYCOSYLATION
CC (DONE IN HB-EGF PURIFIED FROM HISTIOCYTIC LYMPHOMA CELL LINE
CC U-937).
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; M60278; AAA35956.1; -.
DR EMBL; AC004634; AAC15470.1; -.
DR PIR; A38432; A38432.
DR PDB; 1XDT; 25-FEB-98.
DR MIM; 126150; -.
DR InterPro; IPR000561; EGF-like.
DR Pfam; PF00008; EGF; 1.
DR SMART; SM00181; EGF; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
KW Growth factor; Heparin-binding; Signal; EGF-like domain;
KW Transmembrane; Glycoprotein; Receptor; 3D-structure.
FT SIGNAL 1 19
FT PROPEP 20 62
FT CHAIN 63 148
FT PROPEP 149 208
FT DOMAIN 20 160
FT TRANSMEM 161 184
FT DOMAIN 185 208
FT DOMAIN 104 144
FT DISULFID 108 121
FT DISULFID 116 132
FT DISULFID 134 143
FT CARBOHYD 75 75
FT CARBOHYD 85 85
SQ SEQUENCE 208 AA; 23067 MW; 2C43C9D1D8291B51 CRC64;

Query Match 3.1%; Score 7; DB 1; Length 208;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 LPSVVL 10
Db 3 LPSVVL 9

RESULT 11
HBGF_PIG STANDARD; PRT; 208 AA.
AC 001580;
DT 01-FEB-1996 (Rel. 33, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Heparin-binding EGF-like growth factor precursor (HB-EGF) (HBEGF).
GN DTR OR HBGF.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_Taxid=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Pascall J.C.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 99-182 FROM N.A.

```

**us-09-853-526-70.rsp**

QY	db	4 LLP:SVL 10	1111111	3 LLP:SVL 9
Query Match		3.1%;	Score 7;	DB 1;
Best Local Similarity		100.0%;	Pred. No. 12;	
Matches	7;	Conservative	0;	Mismatches
			0;	Indels
			0;	Gaps
			0;	
RESULT	13			

PLSC\_BORBU STANDARD; PRT; 250 AA.

AC 059188;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE 1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51) (1-AGP acyltransferase) (1-AGPAT) (lysophosphatidic acid acyltransferase) (LPAT).  
 GN PLSC OR B80037.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_TaxID=139;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 35210 / H31;  
 RX MEDLINE=98065943; PubMed=9403685;  
 RA Frazer C.M., Casjens S., Huang W.M., Sutton G.G., Clayton F.A., Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M., Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D., Peterson J., Kirlavaga A.R., Quackenbush J., Salzberg S., Hanson M., van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman C., Ulielback T., Matthey L., McDonald L., Artlach P., Bowman C., Gierland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B., Smith H.O., Venter J.C.;  
 RT "Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi."  
 RL Nature 390:580-586(1997).  
 RN [2]  
 RP SEQUENCE OF 116-250 FROM N.A.  
 RC STRAIN=212;  
 RX MEDLINE=95111614; PubMed=7812434;  
 RA Ojalml C., Davidson B.E., Saint-Girons I., Old I.G.;  
 RT "Conservation of gene arrangement and an unusual organization of rRNA genes in the linear chromosomes of the Lyme disease spirochaetes  
 RT Borrelia burgdorferi, B. garinii and B. afzelii."  
 RL Microbiology 140:2931-2940(1994).  
 CC -1- FUNCTION: CONVERTS LYSOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION.  
 CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate -> CoA + 1,2-diacyl-sn-glycerol 3-phosphate.  
 CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.  
 CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AE001117; AAC66417.1; -.  
 DR EMBL: L32861; AAC41407.1; -.  
 DR TIGR: B80037; -.  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KM Phospholipid biosynthesis; transferase; Acyltransferase;  
 KM Complete proteome.  
 SQ SEQUENCE 250 AA; 28580 MW; A730BED7058AD999 CRC64;

Query Match 3.1%; Score 7; DB 1; Length 250;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 159 IPEPGR 165  
 |||||  
 DB 158 IPEPGR 164

RESULT 14

PSD8\_CAEL STANDARD; PRT; 250 AA.

AC 023449;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Probable 26S proteasome regulatory subunit S14.  
 GN ZK20.5.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RA Gajadaty S.;  
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: ACTS AS A REGULATORY SUBUNIT OF THE 26 PROTEASOME WHICH IS INVOLVED IN THE ATP-DEPENDENT DEGRADATION OF UBIQUITINATED PROTEINS (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE PROTEASOME SUBUNIT S14 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: Z69904; CA93778.1; -.  
 DR WormPep: ZK20.5; CE06608.  
 KW Proteasome.  
 SQ SEQUENCE 250 AA; 28795 MW; 7406538A0CE1B32DE CRC64;

Query Match 3.1%; Score 7; DB 1; Length 250;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 171 TKVLSAS 177  
 |||||  
 DB 30 TKVLSAS 36

RESULT 15

PROC\_YEAST STANDARD; PRT; 286 AA.

AC P32263;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Pyruvate-5-carboxylate reductase (EC 1.5.1.2) (P5CR) (P5C reductase).  
 GN PRO3 OR ORE2 OR YER023W.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=92276362; PubMed=1592829;  
 RX MEDLINE=92276362; PubMed=1592829;  
 RA Brandeis M.C., Falvey D.A.;  
 RT "Proline biosynthesis in Saccharomyces cerevisiae: analysis of the RT PRO3 gene, which encodes delta 1-pyrroline-5-carboxylate reductase."  
 RL J. Bacteriol. 174:3782-3788(1992).  
 RN [2]  
 RP ERRATUM.  
 RX MEDLINE=92332464; PubMed=1352771;  
 RA Brandeis M.C., Falvey D.A.;  
 RL J. Bacteriol. 174:5176-5176(1992).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92374987; PubMed=1508147;

RA Neuville P., Aigle M.;  
 RT "ore2, a mutation affecting proline biosynthesis in the yeast  
 RT *Saccharomyces cerevisiae*, leads to a cdc phenotype.";   
 RL Mol. Genet. 234:193-200(1992).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288C / AB972;  
 RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,  
 RA Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,  
 RA Chung E., Duncan M., Guzman E., Hartzell G., Hunnicke-Smith S.,  
 RA Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,  
 RA Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,  
 RA Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,  
 RA Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;  
 RL Submitted (DEC-1994) to the EMBL/Genbank/DDbj databases.  
 CC -1- CATALYTIC ACTIVITY: L-proline + NAD(P)(+) = L-pyrroline-5-  
 CC cardoxylate + NAD(P)H.  
 CC -1- PATHWAY: TERMINAL (THIRD) STEP IN PROLINE BIOSYNTHESIS.  
 CC -1- SUBUNIT: HOMOTETRAMER.  
 CC -1- SIMILARITY: BELONGS TO THE PYRROLINE-5-CARBOXYLATE REDUCTASE  
 CC FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC DR EMBL; M57886; AAA34905.1; -;  
 CC DR EMBL; X57338; CAA40614.1; -;  
 CC DR EMBL; U18778; AAB64556.1; -;  
 CC DR PIR; A41906; A41906.  
 CC DR PIR; A42722; A42722.  
 CC DR PIR; S25293; S25293.  
 CC DR SGD; S0000825; PRO3.  
 CC DR InterPro; IPR000304; PRO3.  
 CC DR Pfam; PF01089; P5CR.1.  
 CC DR PROSITE; PS00521; P5CR.1.  
 CC KM Oxidoreductase; Proline biosynthesis; NADP.  
 CC SQ SEQUENCE 286 AA; 30132 MW; AEB7ID93B46D08B3 CRC64;

Query Match 3 1%; Score 7; DB 1; Length 286;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 7 SYVLIGT 13  
 |||||  
 DB 84 SYVLIGT 90

Search completed: August 28, 2002, 11:27:20  
 Job time: 513 sec

THIS PAGE BLANK (10)

Connecting via Winsock to STN

Welcome to STN International! Enter x:x

LOGINID:ssap1642cy

PASSWORD:  
TERMINAL (ENTER 1, 2, 3, OR ?):2

\*\*\*\*\* Welcome to STN International \*\*\*\*\*

NEWS 1 Web Page URLs for STN Seminar Schedule - N. America  
NEWS 2 "Ask CAS" for self-help around the clock  
NEWS 3 Apr 08 BELSTEIN, Reload and Implementation of a New Subject Area  
NEWS 4 Apr 09 ZDB will be removed from STN  
NEWS 5 Apr 19 US Patent Applications available in IFICDB, IFIPAT, and IFIUDB  
NEWS 6 Apr 22 Records from IP.com available in CAPLUS, HCAPLUS, and ZCAPLUS  
NEWS 7 Apr 22 BIOSIS Gene Names now available in TOXCENTER  
NEWS 8 Apr 22 Federal Research in Progress (FEDRIP) now available  
NEWS 9 Jun 03 New e-mail delivery for search results now available  
NEWS 10 MEDLINE Reload  
NEWS 11 PCTFUL has been reloaded  
NEWS 12 FOREG no longer contains STANDARDS file segment  
NEWS 13 Jul 22 USAN to be reloaded July 28, 2002;  
NEWS 14 saved answer sets no longer valid  
NEWS 15 Jul 29 Enhanced polymer searching in REGISTRY  
NEWS 16 Jul 30 NEFIRST to be removed from STN  
NEWS 17 Aug 08 CANCERLIT reload  
NEWS 18 Aug 08 PHARMAMarketletier (PHARMAML) - new on STN  
NEWS 19 Aug 08 NTIS has been reloaded and enhanced  
NEWS 20 Aug 19 JAPLO to be reloaded August 25, 2002  
NEWS 21 Aug 19 Aquatic Toxicity Information Retrieval (AQUIRE)  
NEWS 22 Aug 19 now available on STN  
NEWS 23 Aug 26 IFIPAT, IFICDB, and IFIUDB have been reloaded  
NEWS 24 The MEDLINE file segment of TOXCENTER has been reloaded  
NEWS 25 Sequence searching in REGISTRY enhanced  
NEWS EXPRESS February 1 CURRENT WINDOWS VERSION IS V6.0d,  
CURRENT MACINTOSH VERSION IS V6.0a(ENG) AND V6.0a(JP),  
AND CURRENT DISCOVER FILE IS DATED 05 FEBRUARY 2002  
NEWS HOURS STN Operating Hours Plus Help Desk Availability  
NEWS INTER General Internet Information  
NEWS LOGIN Welcome Banner and News Items  
NEWS PHONE Direct Dial and Telecommunication Network Access to STN  
NEWS WWW CAS World Wide Web Site (general information)

Enter NEWS followed by the item number or name to see news on that specific topic.

All use of STN is subject to the provisions of the STN Customer agreement. Please note that this agreement limits use to scientific research. Use for software development or design or implementation of commercial gateways or other similar uses is prohibited and may result in loss of user privileges and other penalties.

\*\*\*\*\* STN Columbus \*\*\*\*\*

FILE 'HOME' ENTERED AT 12:44:08 ON 28 AUG 2002

=> file medline, cancerlit, caplus, biosis, confsci, unpatfull, embase  
COST IN U.S. DOLLARS  
FULL ESTIMATED COST  
ENTRY 0.42  
TOTAL SESSION 0.42

FILE 'MEDLINE' ENTERED AT 12:45:03 ON 28 AUG 2002

FILE 'CANCERLIT' ENTERED AT 12:45:03 ON 28 AUG 2002

FILE 'CAPLUS' ENTERED AT 12:45:03 ON 28 AUG 2002  
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.  
PLEASE SEE "HELP USAGETTERMS" FOR DETAILS.

COPYRIGHT (C) 2002 AMERICAN CHEMICAL SOCIETY (ACS)

FILE 'BIOSIS' ENTERED AT 12:45:03 ON 28 AUG 2002  
COPYRIGHT (C) 2002 BIOLOGICAL ABSTRACTS INC. (R)

FILE 'CONFSCI' ENTERED AT 12:45:03 ON 28 AUG 2002  
COPYRIGHT (C) 2002 Cambridge Scientific Abstracts (CSA)

FILE 'USPATFULL' ENTERED AT 12:45:03 ON 28 AUG 2002  
CA INDEXING COPYRIGHT (C) 2002 AMERICAN CHEMICAL SOCIETY (ACS)

FILE 'EMBASE' ENTERED AT 12:45:03 ON 28 AUG 2002  
COPYRIGHT (C) 2002 Elsevier Science B.V. All rights reserved.

=> s PGI  
L1 1251 PGI

=> s l1 and prostate?  
L2 28 l1 AND PROSTATE?

=> d 12 1-28

L2 ANSWER 1 OF 28 MEDLINE  
AN 2001400162 MEDLINE  
DN 21344590 Pubmed ID: 11443539

TI Linkage and association studies of prostate cancer  
susceptibility: evidence for linkage at 8p22-23.

AU Xu J; Zheng S L; Hawkins G A; Faith D A; Kelly B; Isaacs S D; Wiley K E;  
Chang B; Ewing C M; Bujnovsky P; Carpen J D; Blecker E R; Walsh P C;  
Trent J M; Meyers D A; Isaacs W B

CS Center for Human Genomics, Wake Forest University School of Medicine,  
Winston-Salem, NC, USA.

NC CAS8236 (NCI)  
SO AMERICAN JOURNAL OF HUMAN GENETICS, (2001 Aug) 69 (2) 341-50.

CY Journal code: 0370475. ISSN: 0002-9297.  
DT United States  
LA Journal: Article; (JOURNAL ARTICLE)

FS English  
OS Priority Journals  
OMIM-176807; OMIM-300147; OMIM-601518; OMIM-602759; OMIM-603688;  
EM OMIM-605367

ED Entered STN: 20010820  
Last Updated on STN: 20010820  
Entered Medline: 20010816

L2 ANSWER 2 OF 28 CANCERLIT  
AN 2002066870 CANCERLIT  
DN 21344590 Pubmed ID: 11443539  
TI Linkage and association studies of prostate cancer  
susceptibility: evidence for linkage at 8p22-23.

**THIS PAGE BLANK (USPTO)**





**THIS PAGE BLANK (USPTO)**

W: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, GR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LS, LT, LU, LV, LY, MD, MG, MK, MN, MM, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, SM, ST, TU, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, AM, AZ, BY, BG, KZ, MD, RU, TJ, TM, RW: GH, GM, KE, LS, MW, SD, SZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG

US 5945522 A 19990831 US 1997-996306 19971222  
CA 2311416 A 19990701 CA 1998-2311416 19981222  
AU 9915740 A 19990701 AU 1998-15740 19981222  
EP 991770 A 20000412 EP 1998-960061 19981222  
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI

EP 1052292 A1 20001115 EP 2000-115991 19981222  
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI

AT 197817 E 20001215 AT 1998-960061 19981222  
JP 2002516657 T2 20020611 JP 2000-525562 19981222  
US 1997-996306 A 19971222  
US 1998-960061 P 19980909  
WO 1998-182133 W 19981222

L2 ANSWER 8 OF 28 CAPLUS COPYRIGHT 2002 ACS  
AN 1999:81619 CAPLUS  
DN 130:149526  
TI Biallelic markers for use in constructing a high density disequilibrium map of the human genome  
IN Cohen, Daniel; Blumenfeld, Marta; Tchoumakov, Ilya  
PA Genset, Fr.  
SO PCT Int. Appl., 288 pp.  
CODEN: PIXXD2  
DT Patent  
LA English  
FAM, CNT 3

PATENT NO. KIND DATE APPLICATION NO. DATE  
WO 9904038 A2 19990128 WO 1998-181193 19980717  
W: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LS, LT, LU, LV, MD, MG, MK, MN, MM, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, SM, ST, TU, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, AM, AZ, BY, BG, KZ, MD, RU, TJ, TM, RW: GH, GM, KE, LS, MW, SD, SZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG

AU 9884569 A1 19990210 AU 1998-84569 19980717  
AU 746682 B2 20000502 EP 1998-935225 19980717  
EP 1002131 A2 20000524  
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI

PRAI US 1998-82614P 19980421  
EP 1997-401740 A 19970718  
WO 1998-181193 W 19980717

L2 ANSWER 9 OF 28 CAPLUS COPYRIGHT 2002 ACS  
AN 1993:140975 CAPLUS  
DN 118:140975  
TI Rat gene mapping using PCR-analyzed microsatellites  
AU Serikawa, Taduo; Kuramoto, Takashi; Hilbert, Pascale; Mori, Masayuki;

Yamada, Junzo; Dubay, Christopher J.; Lindpaintner, Klaus; Ganten, Detlev; Genset, Jean Louis; et al  
Fac. Med., Kyoto Univ., Kyoto, 606, Japan  
Genetics (1992), 131(3), 701-21  
CODEN: GENTAE; ISSN: 0016-6731

DT Journal  
LA English

L2 ANSWER 10 OF 28 BIOSIS COPYRIGHT 2002 BIOLOGICAL ABSTRACTS INC.  
AN 2002:194891 BIOSIS  
DN PREV200200194891  
TI Prostate cancer gene  
AU Cohen, Daniel (1); Blumenfeld, Marta; Chumakov, Ilya; Bougueleret, Lydie  
CS (1) Fontenay-sous-bois France  
ASSIGNEE: Genset, France  
PI US 6346381 February 12, 2002  
SO Official Gazette of the United States Patent and Trademark Office Patents, (Feb. 12, 2002) Vol. 125, No. 2, pp. No Pagination.  
http://www.uspto.gov/web/menu/patdata.html. e-file.  
ISSN: 0098-1133.  
DT Patent  
LA English

L2 ANSWER 11 OF 28 BIOSIS COPYRIGHT 2002 BIOLOGICAL ABSTRACTS INC.  
AN 2001:432446 BIOSIS  
DN PREV200100432446  
TI Prostate cancer gene.  
AU Cohen, Daniel (1); Blumenfeld, Marta; Chumakov, Ilya; Bougueleret, Lydie  
CS (1) Neuilly sur Seine France  
ASSIGNEE: Genset, France  
PI US 6265546 July 24, 2001  
SO Official Gazette of the United States Patent and Trademark Office Patents, (July 24, 2001) Vol. 1248, No. 4, pp. No Pagination. e-file.  
ISSN: 0098-1133.  
DT Patent  
LA English

L2 ANSWER 12 OF 28 BIOSIS COPYRIGHT 2002 BIOLOGICAL ABSTRACTS INC.  
AN 2001:395125 BIOSIS  
DN PREV200100395125  
TI Linkage and association studies of prostate cancer susceptibility: Evidence for linkage at 8p22-23.  
AU Xu, Jianfeng; Zheng, Signu L.; Hawkins, Gregory A.; Faith, Dennis A.; Kelly, Brian; Isaacs, Sarah D.; Wiley, Kathleen E.; Chang, Bao-li; Swing, Charles W.; Bujnovszky, Pirooska; Carpen, John D.; Blecker, Eugene R.; Walsh, Patrick C.; Trent, Jeffrey M.; Meyers, Deborah A.; Isaacs, William B. (1)  
CS (1) Johns Hopkins Hospital, 600 N. Wolfe street, Marburg 115, Baltimore, MD, 21287; wisacs@hml.edu USA  
SO American Journal of Human Genetics, (August, 2001) Vol. 69, No. 2, pp. 341-350. print.  
ISSN: 0002-9297.  
DT Article  
LA English  
SL English

L2 ANSWER 13 OF 28 BIOSIS COPYRIGHT 2002 BIOLOGICAL ABSTRACTS INC.  
AN 1999:1525274 BIOSIS  
DN PREV1999001525274  
TI Prostate cancer gene.  
AU Cohen, Daniel (1); Chumakov, Ilya; Blumenfeld, Marta; Bougueleret, Lydie  
CS (1) Fontenay-sous-bois France  
ASSIGNEE: GENSET  
PI US 5945522 Aug. 31, 1999

**THIS PAGE BLANK (USPTO)**

SO Official Gazette of the United States Patent and Trademark Office Patents,  
(Aug. 31, 1999) Vol. 1225, No. 5, pp. NO PAGINATION.  
ISSN: 0098-1133.

DI Patent  
LA English  
L2 ANSWER 14 OF 28 USPTAFULL  
AN 2002:157048 USPTAFULL  
TI APOPTOSIS INDUCING MOLECULE II AND METHODS OF USE  
IN EBNER, REINHARD, GAITHERSBURG, MD, UNITED STATES  
YU, GUO-LIANG, BERKELEY, CA, UNITED STATES  
RUBEN, STEVEN M., OLNEY, MD, UNITED STATES  
ZHANG, JUN, BETHESDA, MD, UNITED STATES  
ULBRICH, STEPHEN, ROCKVILLE, MD, UNITED STATES  
ZHAH, YIFAN, GAITHERSBURG, MD, UNITED STATES  
PA Human Genome Sciences (U.S. corporation)  
PI US 2002061647 AI 20020627  
AI US 1999-252656 AI 19990219 (9)  
RLI Continuation-in-part of Ser. No. US 1998-27287, filed on 20 Feb 1998,  
PENDING Continuation-in-part of Ser. No. US 1998-3886, filed on 7 Jan  
1998, ABANDONED Continuation-in-part of Ser. No. US 1997-822953, filed  
on 21 Mar 1997, ABANDONED  
PRAI US 1998-75409P 19980220 (60)  
US 1996-13923P 19960322 (60)  
US 1996-30157P 19961031 (60)  
DT Utility  
FS APPLICATION  
LN.CNT 6195  
INCL INCLM: 435/069.100  
NCL INCLS: 530/350.000; 530/399.000; 514/012.000; 536/023.500  
IC NCLM: 435/069.100  
NCLS: 530/350.000; 530/399.000; 514/012.000; 536/023.500  
[7]  
ICM: A61K038-18  
ICS: C12P021-06; C07H021-04; C07K014-00; G01N033-53  
CAS INDEXING IS AVAILABLE FOR THIS PATENT.  
L2 ANSWER 15 OF 28 USPTAFULL  
AN 2002:126357 USPTAFULL  
TI APOPTOSIS INDUCING MOLECULE II  
IN EBNER, REINHARD, GAITHERSBURG, MD, UNITED STATES  
YU, GUO-LIANG, DARNESTOWN, MD, UNITED STATES  
RUBEN, STEVEN M., OLNEY, MD, UNITED STATES  
ULBRICH, STEPHEN, ROCKVILLE, MD, UNITED STATES  
PA Human Genome Sciences, Inc. (U.S. corporation)  
PI US 2002064869 AI 20020530  
AI US 1998-27287 AI 19980220 (9)  
RLI Continuation-in-part of Ser. No. US 1997-822953, filed on 21 Mar 1997,  
ABANDONED  
PRAI US 1996-30157P 19961031 (60)  
US 1996-13923P 19960322 (60)  
DT Utility  
FS APPLICATION  
LN.CNT 4242  
INCL INCLM: 435/320.100  
NCL INCLS: 435/069.100; 435/325.000; 536/023.500  
NCLM: 435/320.100  
NCLS: 435/069.100; 435/325.000; 536/023.500  
[7]  
ICM: C12N015-63  
ICS: C07H021-04; C12N015-00; C12N015-74; C12N005-06; C12N015-70;  
C12N015-09  
CAS INDEXING IS AVAILABLE FOR THIS PATENT.

L2 ANSWER 16 OF 28 USPTAFULL  
AN 2002:61244 USPTAFULL  
TI Mutated cyclin G1 protein  
IN Gordon, Erlinda Maria, Glendale, CA, UNITED STATES  
PI Hall, Frederick L., Glendale, CA, UNITED STATES  
AI US 2002035079 AI 20020321  
RLI US 2001-796149 AI 20010228 (9)  
DT Continuation-in-part of Ser. No. US 2000-325765, filed on 2 Mar 2000,  
ABANDONED  
FS Utility  
LN.CNT 1822  
INCL INCLM: 514/044.000  
NCL INCLS: 424/094.100  
NCLM: 514/044.000  
NCLS: 424/094.100  
[7]  
ICM: A61K048-00  
ICS: A61K038-43  
CAS INDEXING IS AVAILABLE FOR THIS PATENT.  
L2 ANSWER 17 OF 28 USPTAFULL  
AN 2002:29243 USPTAFULL  
TI Prostate cancer gene  
IN Cohen, Daniel, Fontenay-sous-bois, FRANCE  
Blumenfeld, Marla, Paris, FRANCE  
Chumakov, Ilya, Vaux-le-Penil, FRANCE  
Bougueleret, Lydie, Vanves, FRANCE  
PA Genset, FRANCE (non-U.S. corporation)  
PI US 6346381 BI 19981222 (9)  
AI US 1998-218207 20020212  
RLI Continuation-in-part of Ser. No. US 1997-996306, filed on 22 Dec 1997,  
now Patented  
PRAI US 1998-99658P 19980909 (60)  
DT Utility  
FS GRANTED  
LN.CNT 17612  
INCL INCLM: 435/006.000  
NCL INCLS: 435/091.100; 435/091.200; 536/023.100; 536/024.300  
NCLM: 435/006.000  
NCLS: 435/091.100; 435/091.200; 536/023.100; 536/024.300  
[7]  
ICM: C120001-68  
ICS: C12P019-34; C07M021-02  
EXF 435/6; 435/91.2; 435/91.1; 536/23.1; 536/24.3  
CAS INDEXING IS AVAILABLE FOR THIS PATENT.  
L2 ANSWER 18 OF 28 USPTAFULL  
AN 2001:168248 USPTAFULL  
TI NYC homology region II--associated protein and uses thereof  
IN DePinho, Ronald A., Brookline, MA, United States  
PA Albert Einstein College of Medicine of Yeshiva University, Bronx, NY,  
United States (U.S. corporation)  
PI US 6297368 BI 20011002  
AI US 2000-497779 20000203 (9)  
RLI Continuation-in-part of Ser. No. US 1997-946692, filed on 8 Oct 1997,  
now Patented, Pat. No. US 6040425  
DT Utility  
FS GRANTED  
LN.CNT 514  
INCL INCLM: 536/023.500  
NCL INCLS: 530/350.000; 530/300.000; 530/827.000; 514/012.000  
NCLM: 536/023.500  
NCLS: 530/300.000; 530/350.000; 530/827.000

**THIS PAGE BLANK (USPTO)**

IC [7]  
 ICM: C07H021-04  
 ICS: C07K014-47  
 EXP 536/23.1; 536/23.5; 530/350  
 CAS INDEXING IS AVAILABLE FOR THIS PATENT.

L2 ANSWER 19 OF 28 USPTAFULL  
 AN 2001:158014 USPTAFULL  
 TI Methode, software and apparatus for identifying genomic regions harboring a gene associated with a detectable trait  
 IN Schork, Nicholas J., Shaker Heights, OH, United States  
 Cohen-Akenine, Annick, Paris, France  
 Blumenfeld, Marta, Paris, France  
 Cohen, Daniel, Neuilly-sur-Seine, France  
 Gense, Paris, France (non-U.S. corporation)  
 PI US 6291182 B1 20010918  
 AI US 1999-438016 19991110 (9)  
 PRAI US 1998-107986P 19981110 (60)  
 US 1999-140785P 19990623 (60)  
 DT Utility  
 FS GRANTED  
 LN.CNT 4241  
 INCL INCLM: 435/006.000  
 INCLS: 702/027.000  
 NCL NCLM: 435/006.000  
 NCLS: 702/027.000  
 IC [7]  
 ICM: C120001-68  
 ICS: G01N031-00  
 EXP 435/6; 530/300; 530/324; 530/325; 530/326; 712/200; 702/27  
 CAS INDEXING IS AVAILABLE FOR THIS PATENT.

L2 ANSWER 20 OF 28 USPTAFULL  
 AN 2001:117151 USPTAFULL  
 TI Prostate cancer gene  
 IN Cohen, Daniel, Neuilly sur Seine, France  
 Blumenfeld, Marta, Paris, France  
 Chumakov, Ilya, Vaux-le-Penil, France  
 Bougueleret, Lydie, Vanves, France  
 Gense, France (non-U.S. corporation)  
 PI US 6265546 B1 20010724  
 AI US 1999-338907 19990623 (9)  
 RLI Continuation-in-part of Ser. No. US 1998-218207, filed on 22 Dec 1998, now patented, Pat. No. US 5545522  
 PRAI US 1998-99658P 19980909 (60)  
 US 1998-107986P 19981110 (60)  
 DT Utility  
 FS GRANTED  
 LN.CNT 7782  
 INCL INCLM: 530/350.000  
 INCLS: 424/174.100; 435/007.100; 514/002.000  
 NCL NCLM: 530/350.000  
 NCLS: 424/174.100; 435/007.100  
 IC [7]  
 ICM: C07K001-00  
 ICS: A61K039-395; A01N037-18; G01N035-53  
 EXP 530/350; 530/300; 530/324; 530/325; 514/2; 514/10; 514/12; 424/174.1; 435/7.1  
 CAS INDEXING IS AVAILABLE FOR THIS PATENT.

L2 ANSWER 21 OF 28 USPTAFULL  
 AN 2001:86218 USPTAFULL

TI Sequence specific DNA binding p53  
 IN Vogelstein, Bert, Baltimore, MD, United States  
 Kinzler, Kenneth W., Belair, MD, United States  
 Sherman, Michael I., Allendale, NJ, United States  
 The Johns Hopkins University, Baltimore, MD, United States (U.S. corporation)  
 PA Pharmagenics, Inc., Allendale, NJ, United States (U.S. corporation)  
 PI US 6245515 B1 20010612  
 AI US 1999-399773 19990921 (9)  
 RLI Division of Ser. No. US 1994-299074, filed on 1 Sep 1994, now patented, Pat. No. US 5955263 Division of Ser. No. US 1992-860758, filed on 31 Mar 1992, now patented, Pat. No. US 5362623 Continuation-in-part of Ser. No. US 1991-75182, filed on 14 Jun 1991, now abandoned  
 DT Utility  
 FS GRANTED  
 LN.CNT 1804  
 INCL INCLM: 435/006.000  
 INCLS: 536/023.100; 536/024.330; 536/024.500; 536/025.300; 536/027.000  
 NCL NCLM: 435/006.000  
 NCLS: 536/023.100; 536/024.330; 536/024.500; 536/025.300  
 IC [7]  
 ICM: C120001-68  
 ICS: C07H021-00; C07H021-02; C07H021-04; C07H019-00  
 EXP 435/6; 536/23.1; 536/24.33; 536/24.5; 536/25.3; 536/27  
 CAS INDEXING IS AVAILABLE FOR THIS PATENT.

L2 ANSWER 22 OF 28 USPTAFULL  
 AN 2000:117520 USPTAFULL  
 TI G-protein coupled receptor protein and a DNA encoding the receptor  
 IN Hinuma, Shuji, Tsukuba, Japan  
 Hosoya, Masaki, Tsukuba, Japan  
 Fujii, Ryo, Tsukuba, Japan  
 Ohtaki, Tetsuya, Tsukuba, Japan  
 Fukusumi, Shoji, Tsukuba, Japan  
 Ongi, Kazuhiro, Tsukuba, Japan  
 Takeda Chemical Industries, Ltd., Osaka, Japan (non-U.S. corporation)  
 PI US 6114139 20000905  
 WO 9605302 19960222  
 AI US 1995-513974 19950914 (8)  
 WO 1995-JP1599 19950810  
 PRAI JP 1994-189272 19940811  
 JP 1994-189273 19940811  
 JP 1994-189274 19940811  
 JP 1994-236356 19940930  
 JP 1994-236357 19940930  
 JP 1994-270017 19941102  
 JP 1994-326611 19941228  
 JP 1995-7177 19950120  
 JP 1995-57186 19950316  
 JP 1995-93989 19950419  
 DT Utility  
 FS GRANTED  
 LN.CNT 1308  
 INCL INCLM: 435/069.100  
 INCLS: 530/350.000; 536/023.500; 435/325.000; 435/320.100  
 NCL NCLM: 435/069.100  
 NCLS: 435/320.100; 435/325.000; 530/350.000; 536/023.500  
 IC [7]  
 ICM: C12N015-12  
 ICS: C12N015-63; C12N005-10; C07K014-705  
 EXP 536/23.5; 435/69.1; 435/325; 435/320.1; 530/350  
 CAS INDEXING IS AVAILABLE FOR THIS PATENT.

**THIS PAGE BLANK (USPTO)**



L2 ANSWER 23 OF 28 USPATFUL  
AN 2000:34669 USPATFUL  
TI myc homology region II--associated protein  
IN Depinho, Ronald A., Pelham Manor, NY, United States  
PA Albert Einstein College of Medicine of Yeshiva University, Bronx, NY.  
PI US 6040425  
AI US 1997-946692 20000321  
DT Utility 19971008 (8)  
FS Granted  
LN.CNT 554  
INCL INCLM: 530/350.000  
NCL INCLS: 530/300.000; 530/827.000; 514/012.000  
NCLM: 530/350.000  
NCLS: 530/300.000; 530/827.000  
IC ICM: C07K014-47  
ICS: A61K038-17  
EXF 530/300; 530/350; 530/827; 514/12  
CAS INDEXING IS AVAILABLE FOR THIS PATENT.

L2 ANSWER 24 OF 28 USPATFUL  
AN 1999:113551 USPATFUL  
TI Sequence specific DNA binding by p53  
IN Vogelstein, Bert, Baltimore, MD, United States  
Kinzler, Kenneth W., Baltimore, MD, United States  
Sherman, Michael I., Glen Ridge, NJ, United States  
PA Johns Hopkins University, Baltimore, MD, United States (U.S. corporation)  
PI US 5955263 19990921  
AI US 1994-299074 19940901 (8)  
R1 Division of Ser. No. US 1992-860758, filed on 31 Mar 1992, now patented, Pat. No. US 5362623 which is a continuation-in-part of Ser. No. US 1991-715182, filed on 14 Jun 1991, now abandoned  
DT Utility  
FS Granted  
LN.CNT 2012  
INCL INCLM: 435/006.000  
NCL INCLS: 435/007.230; 436/063.000; 436/064.000; 536/023.100; 536/024.100  
NCLM: 435/006.000  
NCLS: 435/007.230; 436/063.000; 436/064.000; 536/023.100; 536/024.100  
IC ICM: C120001-68  
ICS: G01N033-574; G01N033-48; C07H021-02  
EXF 435/6; 435/7.23; 435/7.92; 536/23.5; 536/24.1; 536/24.1; 436/63; 436/64  
CAS INDEXING IS AVAILABLE FOR THIS PATENT.

L2 ANSWER 25 OF 28 USPATFUL  
AN 1999:102911 USPATFUL  
TI Prostate cancer gene  
IN Cohen, Daniel, Fontenay-sous-bois, France  
Chumakov, Ilya, Vaux-le-Penil, France  
Blumenfeld, Lydie, Vanves, France  
PA Bougueleret, Paris, France  
GENSER, Paris, France (non-U.S. corporation)  
PI US 5945522 19990831  
AI US 1997-996306 19971222 (8)  
DT Utility  
FS Granted  
LN.CNT 5177  
INCL INCLM: 536/023.100  
NCL INCLS: 435/006.000; 536/024.100; 536/024.310; 536/024.320; 536/024.330;

NCL INCLM: 536/024.300  
NCLS: 435/006.000; 536/024.100; 536/024.300; 536/024.310; 536/024.320; 536/024.330  
IC ICM: C07H021-02  
ICS: C07H021-04; C120001-68  
EXF 435/6; 536/24.3; 536/24.31; 536/24.32; 536/24.33; 536/24.1; 536/23.1  
CAS INDEXING IS AVAILABLE FOR THIS PATENT.

L2 ANSWER 26 OF 28 USPATFUL  
AN 97:51869 USPATFUL  
TI Isolated nucleic acid encoding a ubiquitous nuclear receptor  
IN Liao, Shursung, Chicago, IL, United States  
Song, Ching, Durham, NC, United States  
PA Arch Development Corporation, Chicago, IL, United States (U.S. corporation)  
PI US 5639616 19970617  
AI US 1994-342411 19941118 (8)  
R1 Continuation-in-part of Ser. No. US 1993-152003, filed on 10 Nov 1993, now abandoned  
DT Utility  
FS Granted  
LN.CNT 4472  
INCL INCLM: 435/007.100  
NCL INCLS: 435/069.100; 435/252.300; 435/320.100; 536/023.500; 536/024.300  
NCLM: 435/007.100  
NCLS: 435/069.100; 435/252.300; 435/320.100; 536/023.500; 536/024.300  
IC ICM: C12N005-10  
ICS: C12N015-12  
EXF 435/67.1; 435/69.1; 435/252.3; 435/320.1; 536/23.5; 536/24.3  
CAS INDEXING IS AVAILABLE FOR THIS PATENT.

L2 ANSWER 27 OF 28 USPATFUL  
AN 94:97462 USPATFUL  
TI Sequence specific DNA binding by p53  
IN Vogelstein, Bert, Baltimore, MD, United States  
Kinzler, Kenneth W., Baltimore, MD, United States  
Sherman, Michael I., Glen Ridge, NJ, United States  
PA The John Hopkins University, Baltimore, MD, United States (U.S. corporation)  
PI US 5362623 19941108  
AI US 1992-860758 19920331 (7)  
R1 Continuation-in-part of Ser. No. US 1991-715182, filed on 14 Jun 1991  
DT Utility  
FS Granted  
LN.CNT 1741  
INCL INCLM: 435/006.000  
NCL INCLS: 435/320.100; 536/024.100; 536/024.310; 424/002.000  
NCLM: 435/006.000  
NCLS: 435/320.100; 536/024.100; 536/024.310  
IC ICM: C120001-68  
ICS: 435/6; 435/7.1; 435/7.21; 435/7.23; 435/7.92; 435/7.94; 435/70.3; 435/272; 435/975; 435/320.1; 436/501; 436/518; 436/63; 436/64; 436/811; 436/813; 424/2; 536/23.1; 536/24.31; 536/24.1  
EXF 435/6; 435/7.1; 435/7.21; 435/7.23; 435/7.92; 435/7.94; 435/70.3; 435/272; 435/975; 435/320.1; 436/501; 436/518; 436/63; 436/64; 436/811; 436/813; 424/2; 536/23.1; 536/24.31; 536/24.1  
CAS INDEXING IS AVAILABLE FOR THIS PATENT.

L2 ANSWER 28 OF 28 EMBASE COPYRIGHT 2002 ELSEVIER SCI. B.V.  
AN 2001267256 EMBASE  
TI Linkage and association studies of prostate cancer susceptibility. Evidence for linkage at 8p22-23.

**THIS PAGE BLANK (USPTO)**

AU Xu J.; Zheng S.L.; Hawkins G.A.; Faith D.A.; Kelly B.; Isaacs S.D.; Willey K.E.; Chang B.-L.; Ewing C.M.; Bujnovszky P.; Carpten J.D.; Bleecker E.R.; Walsh P.C.; Trent J.M.; Meyers D.A.; Isaacs W.B.  
 CS Dr. W.B. Isaacs, Marburg 115, Johns Hopkins Hospital, 600 N. Wolfe Street, Baltimore, MD 21287, United States. wisacse@jhm.edu  
 SO American Journal of Human Genetics, (2001) 69/2 (341-350).  
 Refs: 36  
 ISSN: 0002-9297 CODEN: AJHGAC  
 CY United States  
 DT Journal; Article  
 FS 016 Cancer  
 022 Human Genetics  
 028 Urology and Nephrology  
 LA English  
 SL English  
 => d 12 1-28d his  
 '1-28d' IS NOT VALID HERE  
 For an explanation, enter "HELP DISPLAY HISTORY".  
 => d his  
 (FILE 'HOME' ENTERED AT 12:44:08 ON 28 AUG 2002)  
 FILE 'MEDLINE, CANCERLIT, CAPLUS, BIOSIS, CONFCL, USPATFULL, EMBASE'  
 ENTERED AT 12:45:03 ON 28 AUG 2002  
 L1 1251 S PGI  
 L2 28 S L1 AND PROSTATE?  
 L3 => s antibod? and 11  
 120 ANTIBOD? AND L1  
 L4 => s 13 and prostate?  
 16 L3 AND PROSTATE?  
 L5 => duprem 14  
 DUPREM IS NOT A RECOGNIZED COMMAND  
 The previous command name entered was not recognized by the system.  
 For a list of commands available to you in the current file, enter  
 "HELP COMMANDS" at an arrow prompt (=>).  
 => dup rem 14  
 PROCESSING COMPLETED FOR L4  
 L5 16 DUP REM L4 (0 DUPLICATES REMOVED)  
 => d 15 1-16  
 L5 ANSWER 1 OF 16 USPATFULL  
 AN 2002:157048 USPATFULL  
 TI APOPTOSIS INDUCING MOLECULE II AND METHODS OF USE  
 IN EBNER, REINHARD, GAITHERSBURG, MD, UNITED STATES  
 YU, GUO-LIANG, BERKELEY, CA, UNITED STATES  
 RUBEN, STEVEN M., OLNEY, MD, UNITED STATES  
 ZHANG, JUN, BETHESDA, MD, UNITED STATES  
 ULBRICH, STEPHEN, ROCKVILLE, MD, UNITED STATES  
 ZHAI, YIPAN, GAITHERSBURG, MD, UNITED STATES  
 Human Genome Sciences (U.S. corporation)  
 PI US 2002061647 AI 20020627  
 P1 US 1999-252656 AI 19990219 (9)  
 RLI Continuation-in-part of Ser. No. US 1998-27287, filed on 20 Feb 1998,  
 PENDING Continuation-in-part of Ser. No. US 1998-3886, filed on 7 Jan  
 1998, ABANDONED Continuation-in-part of Ser. No. US 1997-822953, filed  
 on 21 Mar 1997, ABANDONED

PRAI US 1998-75403P 19980220 (60)  
 US 1996-13923P 19960322 (60)  
 US 1996-30157P 19961031 (60)  
 DT Utility  
 FS APPLICATION  
 LN CNT 6195  
 INCL INCLM: 435/069.100  
 INCL: 530/350.000; 530/399.000; 514/012.000; 536/023.500  
 NCL NCLM: 435/069.100  
 NCLS: 530/350.000; 530/399.000; 514/012.000; 536/023.500  
 IC [7]  
 ICM: A61K038-18  
 ICS: C12P021-06; C07H021-04; C07K014-00; G01N033-53  
 CAS INDEXING IS AVAILABLE FOR THIS PATENT.  
 L5 ANSWER 2 OF 16 USPATFULL  
 AN 2002:126357 USPATFULL  
 TI APOPTOSIS INDUCING MOLECULE II  
 IN EBNER, REINHARD, GAITHERSBURG, MD, UNITED STATES  
 YU, GUO-LIANG, DARNESTOWN, MD, UNITED STATES  
 RUBEN, STEVEN M., OLNEY, MD, UNITED STATES  
 ULBRICH, STEPHEN, ROCKVILLE, MD, UNITED STATES  
 Human Genome Sciences, Inc. (U.S. corporation)  
 PI US 2002064869 AI 20020530  
 P1 US 1998-27287 AI 19980220 (9)  
 RLI Continuation-in-part of Ser. No. US 1997-822953, filed on 21 Mar 1997,  
 ABANDONED  
 PRAI US 1996-30157P 19961031 (60)  
 US 1996-13923P 19960322 (60)  
 DT Utility  
 FS APPLICATION  
 LN CNT 4242  
 INCL INCLM: 435/320.100  
 INCL: 435/069.100; 435/325.000; 536/023.500  
 NCL NCLM: 435/320.100  
 NCLS: 435/069.100; 435/325.000; 536/023.500  
 IC [7]  
 ICM: C12N015-63  
 ICS: C07H021-04; C12N015-00; C12N015-74; C12N005-06; C12N015-70;  
 C12N015-09  
 CAS INDEXING IS AVAILABLE FOR THIS PATENT.  
 L5 ANSWER 3 OF 16 USPATFULL  
 AN 2002:61244 USPATFULL  
 TI Mutated cyclin G1 protein  
 IN Gordon, Erlinda Maria, Glendale, CA, UNITED STATES  
 Hall, Frederick L., Glendale, CA, UNITED STATES  
 PI US 2002035079 AI 20020321  
 P1 US 2001-796149 AI 20010228 (9)  
 RLI Continuation-in-part of Ser. No. US 2000-325765, filed on 2 Mar 2000,  
 ABANDONED  
 DT Utility  
 FS APPLICATION  
 LN CNT 1822  
 INCL INCLM: 514/044.000  
 INCL: 424/094.100  
 NCL NCLM: 514/044.000  
 NCLS: 424/094.100  
 IC [7]  
 ICM: A61K048-00  
 ICS: A61K038-43  
 CAS INDEXING IS AVAILABLE FOR THIS PATENT.  
 L5 ANSWER 4 OF 16 USPATFULL

**THIS PAGE BLANK (USPTO)**

AN 2002:29243 USPTFULT  
 TI Prostate cancer gene  
 IN Cohen, Daniel, Fontenay-sous-bois, FRANCE  
 Blumenfeld, Marta, Paris, FRANCE  
 Chumakov, Ilya, Vaux-le-Penil, FRANCE  
 Bougueleret, Lydie, Vanves, FRANCE  
 PA Genset, FRANCE (non-U.S. corporation)  
 PI US 6346381 B1 20020212  
 AI US 1998-218207 19981222 (9)  
 RLI Continuation-in-part of Ser. No. US 1997-996306, filed on 22 Dec 1997,  
 now patented, Pat. No. US 5945522  
 PRAI US 1998-99658P 19980909 (60)  
 DT Utility  
 FS GRANTED  
 LN CNT 17612  
 INCL INCLM: 435/006.000  
 INCLM: 435/091.100; 435/091.200; 536/023.100; 536/024.300  
 NCLM: 435/006.000  
 NCLM: 435/091.100; 435/091.200; 536/023.100; 536/024.300  
 IC [7]  
 ICM: C12Q001-68  
 EXF ICS: C12P019-34; C07M021-02  
 435/6; 435/91.2; 435/91.1; 536/23.1; 536/24.3  
 CAS INDEXING IS AVAILABLE FOR THIS PATENT.  
 L5 ANSWER 5 OF 16 USPTFULT  
 AN 2001:168248 USPTFULT  
 TI Myc homology region II--associated protein and uses thereof  
 IN Depinho, Ronald A., Brookline, MA, United States  
 PA Albert Einstein College of Medicine of Yeshiva University, Bronx, NY,  
 United States (U.S. corporation)  
 PI US 6297368 B1 20011002  
 AI US 2000-497779 20000203 (9)  
 RLI Continuation-in-part of Ser. No. US 1997-946692, filed on 8 Oct 1997,  
 now patented, Pat. No. US 6040425  
 DT Utility  
 FS GRANTED  
 LN CNT 514  
 INCL INCLM: 536/023.500  
 INCLM: 530/350.000; 530/300.000; 530/827.000; 514/012.000  
 NCLM: 536/023.500  
 NCLM: 530/300.000; 530/350.000; 530/827.000  
 IC [7]  
 ICM: C07H021-04  
 EXF ICS: C07K014-47  
 536/23.1; 536/23.5; 530/350  
 CAS INDEXING IS AVAILABLE FOR THIS PATENT.  
 L5 ANSWER 6 OF 16 USPTFULT  
 AN 2001:18014 USPTFULT  
 TI Methode, software and apparatus for identifying genomic regions harboring  
 a gene associated with a detectable trait  
 IN Schork, Nicholas J., Shaker Heights, OH, United States  
 Essioux, Laurent, Paris, France  
 Cohen-Akenine, Annick, Paris, France  
 Blumenfeld, Marta, Paris, France  
 Cohen, Daniel, Neuilly-sur-Seine, France  
 PA Genset, Paris, France (non-U.S. corporation)  
 PI US 6291182 B1 20010918  
 AI US 1998-438016 19981110 (9)  
 PRAI US 1998-107986P 19981110 (60)  
 DT US 1999-140785P 19990623 (60)  
 FS Utility  
 GRANTED

LN CNT 4241  
 INCL INCLM: 435/006.000  
 INCLM: 702/027.000  
 NCLM: 435/006.000  
 NCLM: 702/027.000  
 IC [7]  
 ICM: C12Q001-68  
 EXF ICS: G01N031-00  
 435/6; 530/300; 530/324; 530/325; 530/326; 712/200; 702/27  
 CAS INDEXING IS AVAILABLE FOR THIS PATENT.  
 L5 ANSWER 7 OF 16 USPTFULT  
 AN 2001:117151 USPTFULT  
 TI Prostate cancer gene  
 IN Cohen, Daniel, Neuilly sur Seine, France  
 Blumenfeld, Marta, Paris, France  
 Chumakov, Ilya, Vaux-le-Penil, France  
 Bougueleret, Lydie, Vanves, France  
 PA Genset, France (non-U.S. corporation)  
 PI US 6265546 B1 20010724  
 AI US 1999-338907 19990623 (9)  
 RLI Continuation-in-part of Ser. No. US 1998-218207, filed on 22 Dec 1998  
 now patented, Pat. No. US 5945522  
 PRAI US 1998-99658P 19980909 (60)  
 DT US 1998-107986P 19981110 (60)  
 FS Utility  
 GRANTED  
 LN CNT 7782  
 INCL INCLM: 530/350.000  
 INCLM: 424/174.100; 435/007.100; 514/002.000  
 NCLM: 530/350.000  
 NCLM: 424/174.100; 435/007.100  
 IC [7]  
 ICM: C07K001-00  
 EXF ICS: A61K039-395; A01N037-18; G01N035-53  
 530/350; 530/300; 530/324; 530/325; 514/2; 514/10; 514/12; 424/174.1;  
 435/7.1  
 CAS INDEXING IS AVAILABLE FOR THIS PATENT.  
 L5 ANSWER 8 OF 16 USPTFULT  
 AN 2001:86218 USPTFULT  
 TI Sequence specific DNA binding p53  
 IN Vogelstein, Bert, Baltimore, MD, United States  
 Kinzler, Kenneth W., Belair, MD, United States  
 Sherman, Michael I., Allendale, NJ, United States  
 PA The Johns Hopkins University, Baltimore, MD, United States (U.S.  
 corporation)  
 PI US 6245515 B1 20010612  
 AI US 1999-399773 19990921 (9)  
 RLI Division of Ser. No. US 1994-299074, filed on 1 Sep 1994, now patented,  
 Pat. No. US 595263, Division of Ser. No. US 1992-860758, filed on 31 Mar  
 1992, now patented, Pat. No. US 536263 Continuation-in-part of Ser. No.  
 US 1991-715182, filed on 14 Jun 1991, now abandoned  
 DT Utility  
 FS GRANTED  
 LN CNT 1804  
 INCL INCLM: 435/006.000  
 INCLM: 435/023.100; 536/024.300; 536/024.500; 536/025.300; 536/027.000  
 NCLM: 435/006.000  
 NCLM: 435/023.100; 536/024.300; 536/024.500; 536/025.300  
 IC [7]  
 ICM: C12Q001-68

**THIS PAGE BLANK (USPTO)**

[illegible]

**THIS PAGE BLANK (USPTO)**



R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT,  
 IE, FI  
 EP 1052292 A1 20001115 EP 2000-115991 19981222  
 R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT,  
 IE, FI  
 AT 197817 E 20001215 AT 1998-960061 19981222  
 JP 2002516657 T2 20020611 JP 2000-525562 19981222  
 US 1997-996306 A 19971222  
 PRAI US 1998-99658P P 19980909  
 EP 1998-960061 A3 19981222  
 WO 1998-1B2133 W 19981222

ANSWER 13 OF 16 USPTAFULL  
 LN 1999:113551 USPTAFULL  
 TI Sequence specific DNA binding by p53  
 IN Vogelstein, Bert, Baltimore, MD, United States  
 Kinzler, Kenneth W., Baltimore, MD, United States  
 Sherman, Michael I., Glen Ridge, NJ, United States  
 Johns Hopkins University, Baltimore, MD, United States (U.S. corporation)  
 PA Genzyme Corporation, Framingham, MA, United States (U.S. corporation)  
 PI US 5955263 19990921  
 AI US 1994-299074 19940901 (8)  
 RLI Division of Ser. No. US 1997-860758, filed on 31 Mar 1992, now patented, Pat. No. US 5362623 which is a continuation-in-part of Ser. No. US 1991-715182, filed on 14 Jun 1991, now abandoned  
 DT Utility  
 FS Granted  
 LN.CNT 2012  
 INCL INCLM: 435/006.000  
 INCLS: 435/007.230; 436/063.000; 436/064.000; 536/023.100; 536/024.100  
 NCL NCLM: 435/006.000  
 IC [6]  
 ICM: C120001-68  
 ICS: G01N033-574, G01N033-48; C07H021-02  
 EXF 435/6; 435/7.23; 435/7.92; 536/23.5; 536/23.1; 536/24.1; 436/63; 436/64  
 CAS INDEXING IS AVAILABLE FOR THIS PATENT.

ANSWER 14 OF 16 USPTAFULL  
 LN 1999:102911 USPTAFULL  
 TI Prostate cancer gene  
 IN Cohen, Daniel, Fontenay-sous-Bois, France  
 Chumakov, Ilya, Vaux-le-Penil, France  
 Blumenfeld, Marta, Paris, France  
 Bougueleret, Lydie, Vanves, France  
 PA CENSET, Paris, France (non-U.S. corporation)  
 PI US 5945522 19990831  
 AI US 1997-996306 19971222 (8)  
 DT Utility  
 FS Granted  
 LN.CNT 5177  
 INCL INCLM: 536/023.100  
 INCLS: 435/006.000; 536/024.100; 536/024.310; 536/024.320; 536/024.330;  
 536/024.300  
 NCL NCLM: 536/023.100  
 IC [6]  
 ICM: C07H021-02  
 ICS: C07H021-04; C120001-68  
 EXF 435/6; 536/24.3; 536/24.31; 536/24.32; 536/24.33; 536/24.1; 536/23.1  
 CAS INDEXING IS AVAILABLE FOR THIS PATENT.

ANSWER 15 OF 16 USPTAFULL  
 LN 97:51869 USPTAFULL  
 TI Isolated nucleic acid encoding a ubiquitous nuclear receptor  
 IN Liao, Shutsung, Chicago, IL, United States  
 Song, Ching, Durham, NC, United States  
 Arch Development Corporation, Chicago, IL, United States (U.S. corporation)  
 PA US 5639616 19970617  
 AI US 1994-342411 19941118 (8)  
 RLI Continuation-in-part of Ser. No. US 1993-152003, filed on 10 Nov 1993, now abandoned  
 DT Utility  
 FS Granted  
 LN.CNT 4472  
 INCL INCLM: 435/007.100  
 INCLS: 435/069.100; 435/252.300; 435/320.100; 536/023.500; 536/024.300  
 NCL NCLM: 435/007.100  
 IC [6]  
 ICM: C12N005-10  
 ICS: C12N015-12  
 EXF 435/67.1; 435/69.1; 435/252.3; 435/320.1; 536/23.5; 536/24.3  
 CAS INDEXING IS AVAILABLE FOR THIS PATENT.

ANSWER 16 OF 16 USPTAFULL  
 LN 94:97462 USPTAFULL  
 TI Sequence specific DNA binding by p53  
 IN Vogelstein, Bert, Baltimore, MD, United States  
 Kinzler, Kenneth W., Baltimore, MD, United States  
 Sherman, Michael I., Glen Ridge, NJ, United States  
 The John Hopkins University, Baltimore, MD, United States (U.S. corporation)  
 PA Pharmacia, Inc., Allendale, NJ, United States (U.S. corporation)  
 PI US 5362623 19941108  
 AI US 1992-860758 19920331 (7)  
 RLI Continuation-in-part of Ser. No. US 1991-715182, filed on 14 Jun 1991  
 DT Utility  
 FS Granted  
 LN.CNT 1741  
 INCL INCLM: 435/006.000  
 INCLS: 435/320.100; 536/024.100; 536/024.310; 424/002.000  
 NCL NCLM: 435/006.000  
 IC [5]  
 ICM: C120001-68  
 ICS: 435/5; 435/6; 435/7.1; 435/7.21; 435/7.23; 435/7.92; 435/7.94; 435/70.3;  
 435/272; 435/975; 435/320.1; 436/501; 436/518; 436/63; 436/64; 436/811;  
 436/813; 424/2; 536/23.1; 536/24.31; 536/24.1  
 EXF 435/5; 435/6; 435/7.1; 435/7.21; 435/7.23; 435/7.92; 435/7.94; 435/70.3;  
 435/272; 435/975; 435/320.1; 436/501; 436/518; 436/63; 436/64; 436/811;  
 436/813; 424/2; 536/23.1; 536/24.31; 536/24.1  
 CAS INDEXING IS AVAILABLE FOR THIS PATENT.

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:28:33 ; Search time 67.07 Seconds  
(without alignments)  
938.872 Million cell updates/sec

Title: US-09-853-526-5  
Perfect score: 364  
Sequence: 1 MLSTLVHHTYMRRLPSVV.....YVNWFWYGTLLGLAWTIKA 364

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 562222 seqs, 172994929 residues

Word size : 0

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :  
1: SPREMBL.19:\*  
2: sp\_archea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_human:\*  
6: sp\_invertebrate:\*  
7: sp\_mammal:\*  
8: sp\_mhc:\*  
9: sp\_organelle:\*  
10: sp\_phage:\*  
11: sp\_plant:\*  
12: sp\_rodent:\*  
13: sp\_virus:\*  
14: sp\_vertebrate:\*  
15: sp\_unclassified:\*  
16: sp\_rvirus:\*  
17: sp\_bacteriap:\*  
18: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	364	100.0	364	4 Q9BQG4	Q9BQG4 homo sapien
2	252	69.2	353	4 Q9NUO2	Q9NUO2 homo sapien
3	32	8.8	354	11 Q9DIB8	Q9DIB8 mus musculu
4	9	2.5	264	16 Q9POE7	Q9POE7 ureaplasma
5	9	2.5	386	5 Q9SRI2	Q9SRI2 caenorhabdi
6	8	2.2	115	16 Q9SE87	Q9SE87 rhizobium 1
7	8	2.2	236	16 Q9A675	Q9A675 caulobacter
8	8	2.2	262	16 Q9A743	Q9A743 chlamydia t
9	8	2.2	282	16 Q9A380	Q9A380 treponema p
10	8	2.2	411	16 Q9A4V3	Q9A4V3 caulobacter
11	8	2.2	430	2 Q54517	Q54517 acinetobact
12	8	2.2	454	3 Q9PSW2	Q9PSW2 neurospora
13	8	2.2	458	16 Q9KFI0	Q9KFI0 bacillus ha
14	8	2.2	459	16 Q31503	Q31503 bacillus su
15	8	2.2	483	10 Q49927	Q49927 pismu sativ
16	8	2.2	555	5 Q961R9	Q961R9 drosophila

17	8	2.2	557	5 Q9VEX8	Q9VEX8 drosophila
18	8	2.2	1820	5 Q9VVG0	Q9VVG0 drosophila
19	7	1.9	54	6 Q28219	Q28219 cercopithe
20	7	1.9	70	5 Q9NE57	Q9NE57 drosophila
21	7	1.9	78	16 Q9B7C0	Q9B7C0 rhizobium 1
22	7	1.9	116	8 Q94TR2	Q94TR2 neoscopelus
23	7	1.9	132	11 Q62370	Q62370 mus musculu
24	7	1.9	137	12 Q9D581	Q9D581 mus musculu
25	7	1.9	141	6 Q28218	Q28218 leucania se
26	7	1.9	143	16 Q98RB9	Q98RB9 cercopithe
27	7	1.9	148	2 Q9RL25	Q9RL25 mycoplasma
28	7	1.9	148	16 Q9CH79	Q9CH79 streptomyce
29	7	1.9	150	2 Q9AGV0	Q9AGV0 lactococcus
30	7	1.9	162	10 Q9LEP0	Q9LEP0 brucella ab
31	7	1.9	163	2 Q9EX02	Q9EX02 arabidopsis
32	7	1.9	170	5 Q9W4S6	Q9W4S6 pseudomonas
33	7	1.9	175	2 Q9EX12	Q9EX12 pasteurella
34	7	1.9	184	11 Q9CWC3	Q9CWC3 aquifex aeo
35	7	1.9	195	16 Q9KCD7	Q9KCD7 streptomyce
36	7	1.9	196	16 Q9LH8	Q9LH8 mus musculu
37	7	1.9	209	16 Q910U7	Q910U7 bacillus ha
38	7	1.9	209	16 Q9CJH4	Q9CJH4 pseudomonas
39	7	1.9	211	16 Q67841	Q67841 pasteurella
40	7	1.9	220	2 Q9EMW7	Q9EMW7 aquifex aeo
41	7	1.9	225	16 Q97NE9	Q97NE9 streptococ
42	7	1.9	230	16 Q9PHZ5	Q9PHZ5 campylobact
43	7	1.9	234	2 Q32330	Q32330 clostridium
44	7	1.9	240	12 Q9WHT2	Q9WHT2 thailand to
45	7	1.9			

## ALIGNMENTS

RESULT 1  
ID Q9BQG4 PRELIMINARY; PRT: 364 AA.  
AC Q9BQG4: 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE HYPOTHETICAL 42.1 KDA PROTEIN (LYSOPHOSPHATIDIC ACID  
DE ACYLTRANSFERASE-EPSILON) (EC 2.3.1.51).  
GN DKFPZP61C222 OR LPAAT-E.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=AMYGDALA;  
RX MEDLINE=21154917; PubMed=11230166;  
RA Wiemann S., Well B., Wellenreuther R., Gassenhuber J., Glasel S.,  
RA Anstorge W., Boecker M., Bloeker H., Bauersachs S., Blum H.,  
RA Lauber J., Duesterhoeft B., Beyer A., Koehler K., Strack N.,  
RA Mewes H.W., Oltewaelder B., Obermaier B., Tampe J., Heubner D.,  
RA Wambutt R., Korn B., Klein M., Poustka A.;  
RA "Towards a Catalog of Human Genes and Proteins: Sequencing and  
RT Analysis of 500 Novel Complete Protein Coding Human cDNAs.";  
RL Genome Res. 11:422-435(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Leung D.W.;  
RT "Cloning and expression of LPAAT-epsilon.";  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBD databases.  
DR EMBL: AL136587; CAB66522.1; -;  
DR EMBL: AF375789; AAK54809.1; -;  
DR InterPro: IPR002123; Acyltransferase.  
DR Pfam: PF01535; Acyltransferase; Acyltransferase.  
DR Hypothetical protein; Transferase; Acyltransferase.  
SO SEQUENCE 364 AA; 42072 MW; 90A0F87FC7C78081 CRC64;

Query Match 130.0%; Score 364; DB 4; Length 364;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLISLVHTYMRVLLPVSLLGTAFTYVLAAGVWRLSFLPARFYQALDDRLCYCYS 60  
 DB 1 MLISLVHTYMRVLLPVSLLGTAFTYVLAAGVWRLSFLPARFYQALDDRLCYCYS 60  
 QY 61 MVEFFENYGVQVLLYGDLPKKNENIYLANHSTVDMIVADILAIROMALGHVRYLK 120  
 DB 61 MVEFFENYGVQVLLYGDLPKKNENIYLANHSTVDMIVADILAIROMALGHVRYLK 120  
 QY 121 EGLWMLPVCYGFAGHGIYVRSKAFNEKEMNKLOSVDAGTDMYLVIFPEGTNYNE 180  
 DB 121 EGLWMLPVCYGFAGHGIYVRSKAFNEKEMNKLOSVDAGTDMYLVIFPEGTNYNE 180  
 QY 181 QTKVLSAQAFAAORGLAVLKHVLPRIKATHVAFCMKNYLDAYDVTVYVEGKDGGQ 240  
 DB 181 QTKVLSAQAFAAORGLAVLKHVLPRIKATHVAFCMKNYLDAYDVTVYVEGKDGGQ 240  
 QY 241 RRSPTWTEFLCKECPKIHIDRIKDDVPEOEHRMRWLHERPEIKDKMLIEFESPD 300  
 DB 241 RRSPTWTEFLCKECPKIHIDRIKDDVPEOEHRMRWLHERPEIKDKMLIEFESPD 300  
 QY 301 PERKRPFGSVNSKISTIKTLPMSLISGLTAGMLMTDAGRKLVTWYITGLGCLMV 360  
 DB 301 PERKRPFGSVNSKISTIKTLPMSLISGLTAGMLMTDAGRKLVTWYITGLGCLMV 360  
 QY 361 TTKA 364  
 DB 361 TTKA 364

RESULT 2  
 Q9NUO2 PRELIMINARY; PRT; 353 AA.  
 AC Q9NUO2;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE CDNA FLJ11210 FIS, CLONE PLAC1007954.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=PLACENTA.  
 RA Isegal T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sugano S., Shiretori A., Sudo H.,  
 RA Magatsuma M., Hosoliri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
 RA Wakabane S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuno Y.,  
 RA Niinomiya K., Iwayanagi T.;  
 RT "MDO human cDNA sequencing project."  
 RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AK002072; BAA92069.1;  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 SQ SEQUENCE 353 AA; 40799 MW; 0A8A14A2361F1B14 CRC64;

Query Match 69.2%; Score 252; DB 4; Length 353;  
 Best Local Similarity 99.7%; Pred. No. 1,4e-250;  
 Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 12 MRLPSVLLCTAPRYVLAAGVWRLSFLPARFYQALDDRLCYCYSWVLEFFENYTG 71  
 DB 1 MRLPSVLLCTAPRYVLAAGVWRLSFLPARFYQALDDRLCYCYSWVLEFFENYTG 60  
 QY 72 VOILLYGDLPRKNENIYLANHSTVDMIVADILAIROMALGHVRYLKEGLKMLPLYGC 131  
 DB 72 VOILLYGDLPRKNENIYLANHSTVDMIVADILAIROMALGHVRYLKEGLKMLPLYGC 131

DB 61 VOILLYGDLPRKNENIYLANHSTVDMIVADILAIROMALGHVRYLKEGLKMLPLYGC 120  
 QY 132 YFAOHGIIYVRSKAFNEKEMNKLOSVDAGTDMYLVIFPEGTNYNEPOTKYLASQA 191  
 DB 121 YFAOHGIIYVRSKAFNEKEMNKLOSVDAGTDMYLVIFPEGTNYNEPOTKYLASQA 180  
 QY 192 AAORGLAVLKHVLPRIKATHVAFCMKNYLDAYDVTVYVEGKDGGQRRSPPTWTEFL 251  
 DB 192 AAORGLAVLKHVLPRIKATHVAFCMKNYLDAYDVTVYVEGKDGGQRRSPPTWTEFL 240  
 QY 252 CKCEPKIHIDRIKDDVPEOEHRMRWLHERPEIKDKMLIEFESSPPERKRPFGS 311  
 DB 241 CKCEPKIHIDRIKDDVPEOEHRMRWLHERPEIKDKMLIEFESSPPERKRPFGS 300  
 QY 312 VNSKLSIKTLPMSLISGLTAGMLMTDAGRKLVTWYITGLGCLMTTKA 364  
 DB 301 VNSKLSIKTLPMSLISGLTAGMLMTDAGRKLVTWYITGLGCLMTTKA 353

RESULT 3  
 Q9DIE8 PRELIMINARY; PRT; 354 AA.  
 AC Q9DIE8;  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DE 1110013A05RIK PROTEIN.  
 DE 1110013A05RIK.  
 GN 1110013A05RIK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=EMBRYO;  
 RX MEDLINE=24085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Felschmann W., Gaasterland T., Gissi C., King B., Kochia H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nakado I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Straubli F., Suzuki R., Tomita M., Wagner L., Mashio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Womberts P.,  
 RA Sasaki H., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmink L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohlski S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 DR EMBL: AK003649; BAB22915.1;  
 DR MGD: MGI:1915880; 1110013A05RIK.  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 SQ SEQUENCE 354 AA; 40943 MW; 447EEF924B91E800 CRC64;

Query Match 8.8%; Score 32; DB 11; Length 354;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-24;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 185 LSASQAFAAORGLAVLKHVLPRIKATHVAFD 216  
 DB 174 LSASQAFAAORGLAVLKHVLPRIKATHVAFD 205

RESULT 4

Q9POE7 PRELIMINARY; PRT; 264 AA.  
 AC Q9POE7  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE.  
 GN PLSC OR U0344.  
 OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).  
 CC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
 CC Mycoplasmataceae; Ureaplasma.  
 OX NCBI\_TaxID=134821;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SEROVAR 3:  
 RX MEDLINE=20500219; PubMed=11048724;  
 RA Glass J.I., Lefkowitz E.J., Glass J.S., Helner C.R., Chen E.Y.,  
 RA Cassell G.H.;  
 RT "The complete sequence of the mucosal pathogen Ureaplasma  
 RT urealyticum.";  
 RL Nature 407:757-762(2000).  
 DR EMBL; AF002131; AAF30753.1; -  
 DR InterPro; IPR002123; Acyltransferase.  
 DR Pfam; PF01553; Acyltransferase; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 264 AA; 30687 MM; DBCE5727A07C74F7 CRC64;

Query Match 2.5%; Score 9; DB 16; Length 264;  
 Best Local Similarity 100.0%; Pred. No. 1.1;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 168 LVIFPEGTR 176  
 Db 156 LVIFPEGTR 164

RESULT 5  
 Q95RI2 PRELIMINARY; PRT; 386 AA.  
 AC Q95RI2  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE HYPOTHETICAL 44.0 KDA PROTEIN.  
 GN F28B3.9.  
 OS Caenorhabditis elegans.  
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 CC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA None;  
 RT "Genome sequence of the nematode C. elegans: a platform for  
 RT investigating biology. The C. elegans Sequencing Consortium.";  
 RL Science 282:2012-2018(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Geisel C., Kramer J., Smith A.;  
 RT "The sequence of C. elegans cosmid F28B3.";  
 RL Submitted (May-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Waterston R.;  
 RT "Direct Submission.";  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF003136; AAK93853.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 386 AA; 43978 MM; ALE3020C201452A6 CRC64;

Query Match 2.5%; Score 9; DB 5; Length 386;  
 Best Local Similarity 100.0%; Pred. No. 1.5;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 222 LDADYDVTV 230  
 Db 230 LDADYDVTV 238

RESULT 6  
 Q98E87 PRELIMINARY; PRT; 115 AA.  
 AC Q98E87  
 DT 01-OCT-2001 (TREMBlrel. 18, Created)  
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)  
 DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)  
 DE MLI4358 PROTEIN.  
 GN MLI4358.  
 OS Rhizobium loti (Mesorhizobium loti).  
 CC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 CC Phyllobacteriaceae; Mesorhizobium.  
 OX NCBI\_TaxID=381;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-MAFR303099;  
 RX MEDLINE=21082930; PubMed=11214968;  
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
 RA Watanabe A., Ideasa K., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,  
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,  
 RA Takeuchi C., Yamada M., Tabata S.;  
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
 RT Mesorhizobium loti.";  
 RL DNA Res. 7:331-338(2000).  
 DR EMBL; AP003004; BAB51033.1; -  
 KW Complete proteome.  
 SQ SEQUENCE 115 AA; 12335 MM; 6AF43D6451F42BE8 CRC64;

Query Match 2.2%; Score 8; DB 16; Length 115;  
 Best Local Similarity 100.0%; Pred. No. 5.6;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 185 LSASQAF 192  
 Db 31 LSASQAF 38

RESULT 7  
 Q9A675 PRELIMINARY; PRT; 236 AA.  
 AC Q9A675  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE.  
 GN CC2219.  
 OS Caulobacter crescentus.  
 CC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
 CC Caulobacter.  
 OX NCBI\_TaxID=69394;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 19089 / CB15;  
 RX MEDLINE=21173698; PubMed=11259647;  
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 RA Eisen J., Heideberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
 RA Deboy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,  
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,  
 RA Uterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,

RA Salberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
 RT "Complete genome sequence of *Caulobacter crescentus*."  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 DR EMBL: AE005893; AAK24190.1; -  
 DR TIGR: CC2219; -  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KW Transferrase; Acyltransferase; Complete proteome.  
 SQ SEQUENCE 236 AA; 26382 MW; 85694D396B8CA1A0 CRC64;

Query Match 2.2%; Score 8; DB 16; Length 236;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 169 VIRPECTR 176  
 DB 146 VIRPECTR 153

RESULT 8  
 084743 PRELIMINARY; PRT; 262 AA.  
 AC 084743;  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE METAL DEPENDENT HYDROLASE.  
 GN YJCJ OR CT738.  
 OS Chlamydia trachomatis.  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCBI\_Taxid=813;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=D/UW-3/CX;  
 RX MEDLINE=99000809; PubMed=9784136;  
 RA Stephens R.S., Kaiman S., Lammel C.J., Fan J., Marathe R., Aravind L.,  
 RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,  
 RA Davis R.W.;  
 RT "Genome sequence of an obligate intracellular pathogen of humans:  
 RT Chlamydia trachomatis."  
 RL Science 282:754-759(1998).  
 DR EMBL: AE001344; AAC68333.1; -  
 KW Hydrolase; Complete proteome.  
 SQ SEQUENCE 262 AA; 29082 MW; E36A6298A23CEFB CRC64;

Query Match 2.2%; Score 8; DB 16; Length 262;  
 Best Local Similarity 100.0%; Pred. No. 12;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 202 HVLTPRIK 209  
 DB 202 HVLTPRIK 209

RESULT 9  
 083380 PRELIMINARY; PRT; 282 AA.  
 AC 083380;  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE LYSOPHOSPHATIDIC ACID ACYLTRANSFERASE, PUTATIVE.  
 GN TP0361.  
 OS Treponema pallidum.  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.  
 OX NCBI\_Taxid=160;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NICHOLS;  
 RX MEDLINE=98332770; PubMed=9665876;  
 RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,

RA Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,  
 RA Sodergren E., Harsham J.M., McLeod M.P., Salberg S., Peterson J.,  
 RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Uterback T.,  
 RA McDonald L., Artlich P., Bowman C., Cotton M.D., Fujii C., Garland S.,  
 RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,  
 RA Venter J.C.;  
 RT "Complete genome sequence of *Treponema pallidum*, the syphilis  
 RT spirochete."  
 RL Science 281:375-388(1998).  
 DR EMBL: AE001215; AAC65346.1; -  
 DR TIGR: TP0361; -  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KW Transferrase; Acyltransferase; Complete proteome.  
 SQ SEQUENCE 282 AA; 31700 MW; DEF69003CD6C1AD CRC64;

Query Match 2.2%; Score 8; DB 16; Length 282;  
 Best Local Similarity 100.0%; Pred. No. 12;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 169 VIRPECTR 176  
 DB 170 VIRPECTR 177

RESULT 10  
 0944V3 PRELIMINARY; PRT; 411 AA.  
 AC 0944V3;  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE METAL ION EFFLUX MEMBRANE FUSION PROTEIN FAMILY.  
 GN CC2722.  
 OS *Caulobacter crescentus*.  
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
 OC Caulobacter.  
 OX NCBI\_Taxid=69394;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 19089 / CB15;  
 RX MEDLINE=21173698; PubMed=11259647;  
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwin M.L., Haft D.H.,  
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,  
 RA Uterback T., Tran K., Wolf A., Yamathavan J., Ermolaeva M., White O.,  
 RA Salberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
 RT "Complete genome sequence of *Caulobacter crescentus*."  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 DR EMBL: AE005938; AAK24687.1; -  
 DR TIGR: CC2722; -  
 DR InterPro: IPR002215; HlyD.  
 DR Pfam: PF00529; HlyD; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 411 AA; 41818 MW; 1BA7849B9BCFC22E CRC64;

Query Match 2.2%; Score 8; DB 16; Length 411;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 327 ILSGTAG 334  
 DB 380 ILSGTAG 387

RESULT 11  
 054517 PRELIMINARY; PRT; 430 AA.  
 AC 054517;

DT 01-JUN-1998 (TREMblrel. 06, Created)  
DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)  
DT 01-OCT-2001 (TREMblrel. 18, Last annotation update)  
DE 3-DEOXY-D-MANNO-2-OCTULOSONATE TRANSFERASE.  
GN WAAA.  
OS Acinetobacter baumannii.  
OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;  
OC Acinetobacter.  
OX NCBI\_TaxID=470;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 19606; AND ATCC 15308;  
RA Bode C.E., Brabetz W., Brade H.;  
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: 296926; CAB09652.1; -;  
KW EMBL: 296925; CAB09651.1; -;  
SQ TRANSFERASE.  
KW  
SQ SEQUENCE 430 AA; 49070 MW; 695293748185D272 CRC64;

Query Match 2.2%; Score 8; DB 2; Length 430;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 329 SGLTAGML 336  
Db 171 SGLTAGML 178

RESULT 12  
Q9P5W2 PRELIMINARY; PRT; 454 AA.  
AC Q9P5W2;  
DT 01-OCT-2000 (TREMblrel. 15, Created)  
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)  
DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
DE RELATED TO GUANINE DEAMINASE.  
GN B3E4.200.  
OS Neurospora crassa.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Sordariales; Sordariaceae; Neurospora.  
OX NCBI\_TaxID=5141;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,  
RA Nyakatura G., Mewes H.W., Mannhaupt G.;  
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
RX [2]  
RP SEQUENCE FROM N.A.  
RA German Neurospora genome project;  
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AL355931; CAB91408.2; -;  
SQ SEQUENCE 454 AA; 49778 MW; 8C1EC14BAEC421D CRC64;

Query Match 2.2%; Score 8; DB 3; Length 454;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 200 LKHVLP 207  
Db 208 LKHVLP 215

RESULT 13  
Q9KF10 PRELIMINARY; PRT; 458 AA.  
AC Q9KF10;  
DT 01-OCT-2000 (TREMblrel. 15, Created)  
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
DE RNA METHYLTRANSFERASE.  
GN BH0687.

OS Bacillus halodurans.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Bacillus.  
OX NCBI\_TaxID=86665;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-C-125 / JCM 9153;  
RX MEDLINE-20512582; PubMed-11058132;  
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
RA Fujii F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,  
RA Horikoshi K.;  
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus  
RT halodurans and genomic sequence comparison with Bacillus subtilis.";  
RL Nucleic Acids Res. 28:4317-4331(2000).  
DR EMBL: AP001509; BAB04406.1; -;  
DR InterPro: IPR000051; SAM\_bind.  
DR InterPro: IPR001566; TRMA\_1.  
DR PROSITE: PS01230; TRMA\_1; UNKNOWN\_1.  
KW Transferase; Methyltransferase; Complete proteome.  
SQ SEQUENCE 458 AA; 51555 MW; 3D33A661E8891240 CRC64;

Query Match 2.2%; Score 8; DB 16; Length 458;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 178 NPEOTKVL 185  
Db 291 NPEOTKVL 298

RESULT 14  
O31503 PRELIMINARY; PRT; 459 AA.  
AC O31503;  
DT 01-JAN-1998 (TREMblrel. 05, Created)  
DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)  
DT 01-OCT-2001 (TREMblrel. 18, Last annotation update)  
DE YEFA PROTEIN.  
GN YEFA.  
OS Bacillus subtilis.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Bacillus.  
OX NCBI\_TaxID=1423;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE-98044033; PubMed-9384377;  
RC STRAIN-168;  
RA Kunst F., Ogasawara N., Mosser I., Albertini A.M., Alloni G.,  
RA Azevedo V., Bertoletti M.G., Bessieres P., Bolotin A., Borchert S.,  
RA Bortiss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
RA Brouillet S., Brusch C.V., Caldwell B., Capiano V., Carter N.M.,  
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,  
RA Britian K.D., Errington J., Fabret C., Ferrari E., Fougere D.,  
RA Fritsch C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
RA Chim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,  
RA Giusseppl G., Guy B.J., Haga K., Haeleth J., Harwood C.R., Henaut A.,  
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,  
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,  
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,  
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,  
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,  
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
RA Sekiyuchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,  
RA Sorokin A., Taccioni E., Takagi T., Takahashi H., Takemaru K.,  
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,  
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,

RA Viart A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,  
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,  
 RT "The complete genome sequence of the gram-positive bacterium *Bacillus*  
 RT *subtilis*.";  
 RL Mature 390:249-256(1997).

RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.,  
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: Z99107; CAB12493.1; -  
 DR InterPro: IPR002792; DJF90.  
 DR InterPro: IPR000051; SAM\_bind.  
 DR Pfam: PF01938; TRAM\_1.  
 DR PROSITE: PS01230; TRAM\_1; UNKNOWN\_1.  
 DR PROSITE: PS01231; TRAM\_2; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 459 AA; 51819 MW; 35E99547779421D5 CRC64;

Query Match 2.2%: Score 8; DB 16; Length 459;  
 Best Local Similarity 100.0%: Pred. No. 19;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 178 NPEQTKVL 185  
 |||||  
 DB 292 NPEQTKVL 299

RESULT 15  
 O49927  
 ID O49927 PRELIMINARY; PRT; 483 AA.  
 AC O49927;  
 DT 01-JUN-1998 (TREMBlrel. 06, Created)  
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE P54 PROTEIN.  
 GN P54.  
 OS Pisum sativum (Garden pea).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.  
 OX NCBI\_TaxID=3888;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Castillo J., Marquez J.A., Franco L., Ballestar E., Rodrigo M.I.,  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: Y11207; CAA72090.1; -  
 DR HSSP: P50477; ICAU.  
 DR InterPro: IPR001113; Seedstore\_7s.  
 DR Pfam: PF00546; Seedstore\_7s; 1.  
 DR Pfam: PF02808; Seedstore\_7s\_C; 1.  
 DR PF02808; Seedstore\_7s\_C; 1.  
 SQ SEQUENCE 483 AA; 54662 MW; 8127BDAAA0178F3D CRC64;

Query Match 2.2%: Score 8; DB 10; Length 483;  
 Best Local Similarity 100.0%: Pred. No. 20;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 162 AGTPMYLV 169  
 |||||  
 DB 184 AGTPMYLV 191

Search completed: August 28, 2002, 11:28:34  
 Job time: 547 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:07:33 ; Search time 75.55 Seconds  
(without alignments)  
535.154 Million cell updates/sec

Title: US-09-853-526-5

Perfect score: 1924

Sequence: 1 MLSLVLTHTYSMRLLPSPV.....YVNMIVYGTLLGLMTVIKA 364

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq.032802:\*

1: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:\*  
2: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:\*  
3: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:\*  
4: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:\*  
5: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:\*  
6: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:\*  
7: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:\*  
8: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:\*  
9: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:\*  
10: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT:\*  
11: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:\*  
12: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:\*  
13: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:\*  
14: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:\*  
15: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:\*  
16: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:\*  
17: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:\*  
18: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:\*  
19: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:\*  
20: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:\*  
21: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:\*  
22: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1924	100.0	364	20	AAV36729	Human PGI protein
2	1924	100.0	364	22	AAAM0128	Human acyltransferase
3	1924	100.0	364	22	AAU00665	Human lysophosphatidyl transferase
4	1913	99.4	372	22	AAAM41377	Human polypeptide
5	1871	97.2	353	22	AAAB85532	Human secreted protein
6	1871	97.2	353	22	AAAB85552	Human secreted protein
7	1868	97.1	353	22	AAAB93595	Human protein sequence
8	1867	97.0	352	20	AAV36728	Human PGI protein
9	1861	96.7	353	22	AAAM39591	Human polypeptide
10	1644	85.4	315	20	AAV36751	PGI splice variant
11	1568.5	81.5	354	20	AAV36741	Mouse PGI protein

12	1555.5	80.8	300	20	AAV36752	PGI splice variant
13	1498	77.9	291	20	AAV36744	PGI splice variant
14	1427	74.2	259	22	AAAB09653	Human gene 15 enco
15	1325	68.9	261	20	AAV36745	PGI splice variant
16	1203	62.5	228	20	AAV36740	Truncated PGI prot
17	1182.5	61.5	238	20	AAV36743	PGI splice variant
18	976	50.7	185	20	AAV36753	PGI splice variant
19	820	42.6	182	20	AAV36750	PGI splice variant
20	449	23.3	96	20	AAV36749	PGI splice variant
21	401.5	20.9	378	21	AAV36740	Arabidopsis thalia
22	396.5	20.6	352	21	AAV36741	Arabidopsis thalia
23	393.5	20.5	341	21	AAV36742	Arabidopsis thalia
24	390.5	20.3	375	21	AAV36743	Arabidopsis thalia
25	390.5	20.3	375	21	AAV36742	Arabidopsis thalia
26	387.5	20.1	351	21	AAV36735	Arabidopsis thalia
27	387.5	20.1	375	21	AAV36741	Arabidopsis thalia
28	387	20.1	257	22	AAV36740	Novel human diagno
29	387	20.1	980	22	AAV36741	Novel human diagno
30	384.5	20.0	351	21	AAV36742	Arabidopsis thalia
31	381	19.8	343	21	AAV36736	Arabidopsis thalia
32	381	19.8	343	21	AAV36743	Arabidopsis thalia
33	332	17.3	77	20	AAV36742	PGI splice variant
34	330	17.2	66	20	AAV36748	PGI splice variant
35	329	17.1	68	20	AAV36747	PGI splice variant
36	329	17.1	90	20	AAV36746	PGI splice variant
37	328	17.0	374	22	AAU00667	Maize lysophosphat
38	326.5	17.0	374	21	AAU00667	Maize 2-acyltransf
39	326	16.9	374	15	AAV59712	Maize 2-acyltransf
40	317.5	16.5	374	21	AAV59712	Maize 2-acyltransf
41	311.5	16.2	387	21	AAV59482	Jojoba acyltransfe
42	310	16.1	377	17	AAV59482	Limnathes CDNA cl
43	309.5	16.1	294	21	AAV59441	Zea mays protein f
44	308.5	16.0	389	21	AAV59441	Arabidopsis thalia
45	307.5	16.0	389	21	AAV59441	Arabidopsis thalia

ALIGNMENTS

RESULT 1

AAV36729 standard; Protein: 364 AA.

ID AAV36729:

AC AAV36729:

DT 27-SEP-1999 (first entry)

DE Human PGI protein sequence.

KW PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker; cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.

OS Homo sapiens.

PN WO9332644-A2.

PD 01-JUL-1999.

PE 22-DEC-1998; 98MO-IB02133.

PR 09-SEP-1998; 98US-0099658.

PR 22-DEC-1997; 97US-0096306.

XX (GENSET) GENSET.

PI Blumenfeld M, Bougueleret L, Chumakov I, Cohen D; WPI: 1999-405178/34.

DR N-PSDB; AA200870.

PT Use of a prostate cancer associated gene and biallelic markers derived from it

PS Claim 7; Page 190-191; 385bp; English.

XX The invention relates to a mammalian PGI gene and protein, and a set of  
CC PGI allelic markers. The PGI polynucleotide and allelic markers are  
CC used in a hybridisation assay, a sequencing assay, or in an  
CC allele-specific amplification assay for determining the identity of a  
CC nucleotide at a PGI-related allelic marker. The methods can be used to  
CC detect and to assess the risk of developing cancer or prostate cancer.  
CC Early-stage diagnosis of prostate cancer relies on prostate specific  
CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
CC to its inability to discriminate between malignant and non-malignant  
CC affections of the organ. A need exists for both a reliable diagnostic  
CC procedure which would enable early-stage diagnosis, and for preventative  
CC and curative treatments of the disease. The PGI gene can be used for  
CC detection of prostate cancer, and the risk of developing it in the  
CC future, and can also be used to determine therapies for the disease.

XX  
XX Sequence 364 AA;

SO  
Sequence 364 AA;

Query Match 100.0%; Score 1924; DB 20; Length 364;  
Best Local Similarity 100.0%; Pred. No. 5,8e-195;  
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLSLVHTTYSMKRVLPEVVLGTAPTTVLAWGYWRLLSATLPARFYALDDRLCYQOS 60  
DB 1 MLSLVHTTYSMKRVLPEVVLGTAPTTVLAWGYWRLLSATLPARFYALDDRLCYQOS 60  
QY 61 MVLFFENYGVQILLYGDLDPKKNENITYLANHOSTVDWIVADILAIKONALGHVRYLK 120  
DB 61 MVLFFENYGVQILLYGDLDPKKNENITYLANHGSVGVADILAIKONALGHVRYLK 120  
QY 121 EGLMPLPGCYFQOHGSIYVKRSKAKFNEKEMRNKLQSYVDAGTPMYLVIRPEGTRNPE 180  
DB 121 EGLMPLPGCYFQOHGSIYVKRSKAKFNEKEMRNKLQSYVDAGTPMYLVIRPEGTRNPE 180  
QY 121 egllwlpjygyfaghgqilyvkrasakfnekemrnklqsyvdagtpmylvirpegrtynpe 180  
DB 181 QTKVLSAQAAOAGLAVLKHVLTPTKATVAFDCCKNTLDIYDVTYVEGKDGCGQ 240  
DB 181 qtkvlsasqaaaglavlkhvltprtkaavafdccknlydalydvtyvegkddqgq 240  
QY 241 RRESPMTTEFLCKECPKIHIDRIKKDVPEDQEHMRMRWHERPEIKDKMLIEFEYSPD 300  
DB 241 rrespmtteflckepkhihidridkkdvpdeqemrtrwhefeklkdmliefespd 300  
QY 301 PERRRPEGKSVNSKLSIKTLPMSMLISGLTAGMLMTDAGRKLVTNTWYGTLLGCLMW 360  
DB 301 perrrfpgkvsnaklsiktlpsmlilsgltagmlmtdagrklvntwlygtllgclmw 360  
QY 361 TTKA 364  
DB 361 tika 364

RESULT 2  
AAM50128  
ID AAM50128 standard; Protein: 364 AA.  
XX  
XX AAM50128;  
AC  
XX  
XX 21-Dec-2001 (first entry)  
DT  
XX  
XX Human acyltransferase 27417.  
DE  
XX  
XX Acyltransferase 27417; human; cell proliferation; cell migration;  
KM cell differentiation; ovary cancer; brain cancer; colon cancer;  
KM lung cancer; tumour; metastasis; sarcoma; carcinoma;  
KM adenocarcinoma; antitumour; lipid metabolism; diagnosis; therapy.  
XX  
XX Homo sapiens.  
OS  
XX  
XX Key Location/Qualifiers  
FH Peptide 1..49  
FT  
FT /label= Signal-peptide  
?

FT Protein 50..364  
FT /label= Mature\_protein  
FT Domain 50..320  
FT /label= N-terminal\_domain  
FT Domain 321..337  
FT /label= Transmembrane\_domain  
FT Domain 338..364  
FT /label= C-terminal\_domain  
FT Domain 71..363  
FT /label= Acyltransferase\_domain  
FT Modified-site 68..71  
FT /note= "Asn is N-glycosylated"  
FT Modified-site 241..244  
FT /note= "O-phosphorylated by cAMP- and cGMP-dependent protein kinase"  
FT Modified-site 11..13  
FT /note= "O-phosphorylated by protein kinase C"  
FT Modified-site 144..146  
FT /note= "O-phosphorylated by protein kinase C"  
FT Modified-site 205..207  
FT /note= "O-phosphorylated by protein kinase C"  
FT Modified-site 317..319  
FT /note= "O-phosphorylated by protein kinase C"  
FT Modified-site 361..363  
FT /note= "O-phosphorylated by protein kinase C"  
FT Modified-site 95..98  
FT /note= "O-phosphorylated by casein kinase II"  
FT Modified-site 158..161  
FT /note= "O-phosphorylated by casein kinase II"  
FT Modified-site 246..249  
FT /note= "O-phosphorylated by casein kinase II"  
FT Modified-site 82..89  
FT /note= "O-phosphorylated by tyrosine kinase"  
FT Modified-site 219..226  
FT /note= "O-phosphorylated by tyrosine kinase"  
FT Modified-site 23..28  
FT /note= "N-myristoylated"  
FT Modified-site 130..135  
FT /note= "N-myristoylated"  
FT Modified-site 330..335  
FT /note= "N-myristoylated"  
FT Modified-site 352..357  
FT /note= "N-myristoylated"  
FT Modified-site 230..343  
FT /note= "Amidation"  
FT Peptide 105..126  
FT /note= "predicted leucine zipper"  
XX  
XX W0200173051-A2.  
XX  
XX 04-OCT-2001.  
XX  
XX 26-MAR-2001; 2001WO-US09633.  
XX  
XX 24-MAR-2000; 2000US-192092P.  
XX  
XX (MILL-) MILLENNIUM PHARM INC.  
XX  
XX Meyers RA, Rudolph-Owen L, Macbeth KJ, Tsai F, Williamson M;  
XX WPI; 2001-626267/72.  
XX N-PSDB; AAH27094.  
XX  
XX New polynucleotide encodes a polypeptide for the diagnosis and  
XX treatment of cancer and to alter cellular lipid metabolism comprises a  
XX polynucleotide encoding human acyltransferases designated 46743 and  
XX 27417 -  
XX  
XX Claim 9; Fig 8A-B; 150bp; English.  
XX  
XX The present sequence is that of a novel human protein, termed  
XX 27417, which shows the structural characteristics of members of the  
XX acyltransferase family, including an acyltransferase domain. The

CC invention provides novel acyltransferase 27417 nucleic acids and  
CC polypeptides, as well as methods for detecting their presence, and  
CC methods for screening for compounds that modulate their expression  
CC or activity. Such compounds can be used to treat conditions  
CC related to aberrant activity or expression of 27417 protein or  
CC nucleic acid, such as conditions involving deficient cellular  
CC proliferation, migration and/or differentiation. Also provided are  
CC methods for inhibiting the proliferation or migration, or inducing  
CC the killing, of a 27417-expressing cell, e.g. a hyperproliferative  
CC and/or metastatic cell, by contact with a compound that modulates  
CC the activity or expression of the 27417 protein or nucleic acid.  
CC The 27417-expressing cell is found in the healthy or diseased  
CC heart, blood vessels, kidney, skeletal muscle, brain or liver, or  
CC especially in a solid tumour, a soft tissue tumour or a metastatic  
CC lesion, a sarcoma, a carcinoma or an adenocarcinoma, and in  
CC particular in a hyperproliferative and/or metastatic cell found in  
CC ovarian, brain, colon or lung cancer. The compounds may also be  
CC used to modulate lipid metabolism in a 27417-expressing cell.  
CC Methods of disease diagnosis, e.g. by determining the  
CC presence of a genetic alteration in a 27417 polypeptide, and for  
CC evaluating the efficacy of a treatment of a disorder, are also  
CC provided.

XX Sequence 364 AA:

Query Match 100.0%: Score 1924; DB 22; Length 364;  
Best Local Similarity 100.0%: Pred. No. 5.8e-195;  
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MLSTVLHTYSMRKYLPSVLLGTAPTYVLANGVWRLSAPLPARYQALDDRLCYVOS 60  
DB 1 mlstvlhtysmrtyllpsvlllgtapttylavgvwrllsalfparfyqalddrlycvygs 60  
OY 61 MYLFFPENTGVQILLYGDLPRKKNENIYLANHGSTVDMIVADIIARONALGHVRYLK 120  
DB 61 mylffentgyvqillygdlprkknennylnhgstvdmivadiiairgnalghvrylyk 120  
OY 121 EGLKMLPLTGYCYFAOHGCIYVRSKAFNEKEMRNKLOSVDAGTPMYLVIPEGRPYNE 180  
DB 121 eglkmlpltygcyfaohgciyvrskafnekemrnklosvdaagtpmylvipegrtynpe 180  
OY 181 QTKVLSASQAFAORGLAVLKHVLPRIKATVAFDCKMKNYDAIYDVYVVEGKDDGQ 240  
DB 181 qtkvlsasqafaarglavlkhlvprikathvafdcmknydaidydvvyvegkddgq 240  
OY 241 RRESPTMTEFLCKECPKIHIDRIDKDVPEQEHMRRLHREFEIKDKMLIEFYESP 300  
DB 241 rresptmteflckeckpkihidridkdvpeeqehmrwlherfeikdkmliefespd 300  
OY 301 PERRRRFPCKSVNSKLSIKKTLPSMLILSGTRAGMLMTPAGRKLYNWTWYIGTLGCLMV 360  
DB 301 perrrrfpcksvnskslisktlpsmllsgltagmlmtaagrklyvnwtwyigtlgclmv 360  
OY 361 TIKA 364  
DB 361 tika 364

RESULT 3

AAU00665 AAU00665 standard; Protein; 364 AA.

XX AAU00665;

XX 07-SEP-2001 (first entry)

DE Human lysophosphatidic acid acyltransferase isoform LPAAT-epsilon.

XX Lysophosphatidic acid acyltransferase epsilon; LPAAT-epsilon; LPA; PA;  
KM lysophosphatidic acid; phosphatidic acid; acylation; cellular activation;  
KM phospholipid signalling; mitogenesis; inflammation; autoimmune disease;  
KM oncology; cancer; obesity; gene therapy.

XX Homo sapiens.  
OS  
XX WO200134782-A1.  
PN  
XX 17-MAY-2001.  
PD  
XX 02-NOV-2000; 2000MO-US30193.  
PF  
XX 09-NOV-1999; 99US-0436919.  
PR  
XX (CELL-) CELL THERAPEUTICS INC.  
PA  
PI Leung DW;  
XX WPI: 2001-335920/35.  
DR N-PSDB: AAS00649.  
XX  
PT Novel isolated human isoform of lysophosphatidic acid  
PT acyltransferase-epsilon useful for diagnostic, therapeutic and  
PT screening purposes -  
XX  
PS Claim 1; Fig 1; 48pp; English.

The sequence represents a human lysophosphatidic acid acyltransferase (LPAAT) isoform, LPAAT-epsilon. LPAAT catalyses the acylation of lysophosphatidic acid (LPA) to phosphatidic acid (PA). LPA and PA have been identified as phospholipid signalling molecules that affect a wide range of biological responses. PA is involved in cellular activation and mitogenesis. Compounds that block PA generation and hence diminish lipid biosynthesis and the signal involved in cell activation are of therapeutic interest in the areas of inflammation and oncology (e.g. autoimmune diseases and cancer) as well as obesity treatment. LPAAT-epsilon and its corresponding DNA can be used in screening assays to detect agents that stimulate or inhibit the activity of LPAAT and, therefore, PA. The DNA is useful in tests to detect the presence or expression of LPAAT-epsilon in relation to certain diseases and conditions, and in disease prevention and treatment. The sequences of the invention are also useful for diagnosis of diseases and conditions in which the expression of LPAAT enzyme is abnormal.

XX Sequence 364 AA:

Query Match 100.0%: Score 1924; DB 22; Length 364;  
Best Local Similarity 100.0%: Pred. No. 5.8e-195;  
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MLSTVLHTYSMRKYLPSVLLGTAPTYVLANGVWRLSAPLPARYQALDDRLCYVOS 60  
DB 1 mlstvlhtysmrtyllpsvlllgtapttylavgvwrllsalfparfyqalddrlycvygs 60  
OY 61 MYLFFPENTGVQILLYGDLPRKKNENIYLANHGSTVDMIVADIIARONALGHVRYLK 120  
DB 61 mylffentgyvqillygdlprkknennylnhgstvdmivadiiairgnalghvrylyk 120  
OY 121 EGLKMLPLTGYCYFAOHGCIYVRSKAFNEKEMRNKLOSVDAGTPMYLVIPEGRPYNE 180  
DB 121 eglkmlpltygcyfaohgciyvrskafnekemrnklosvdaagtpmylvipegrtynpe 180  
OY 181 QTKVLSASQAFAORGLAVLKHVLPRIKATVAFDCKMKNYDAIYDVYVVEGKDDGQ 240  
DB 181 qtkvlsasqafaarglavlkhlvprikathvafdcmknydaidydvvyvegkddgq 240  
OY 241 RRESPTMTEFLCKECPKIHIDRIDKDVPEQEHMRRLHREFEIKDKMLIEFYESP 300  
DB 241 rresptmteflckeckpkihidridkdvpeeqehmrwlherfeikdkmliefespd 300  
OY 301 PERRRRFPCKSVNSKLSIKKTLPSMLILSGTRAGMLMTPAGRKLYNWTWYIGTLGCLMV 360  
DB 301 perrrrfpcksvnskslisktlpsmllsgltagmlmtaagrklyvnwtwyigtlgclmv 360  
OY 361 TIKA 364

Db 361 tika 364

## RESULT 4

AA041377 standard; protein: 372 AA.

AA041377:

22-OCT-2001 (first entry)

Human polypeptide SEQ ID NO 6308.

Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;

peripheral nervous system; neuropathy; central nervous system; CNS;

Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;

amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;

chemokine; thrombolytic; drug screening; arthritis; inflammation;

Leukemia.

Homo sapiens.

MO200153312-A1.

26-JUL-2001.

26-DEC-2000; 2000MO-US34263.

25-APR-2000; 2000US-0486725.

09-JUL-2000; 2000US-0552317.

19-JUL-2000; 2000US-0598042.

03-AUG-2000; 2000US-0620312.

14-SEP-2000; 2000US-0653450.

19-OCT-2000; 2000US-0662191.

29-NOV-2000; 2000US-0693036.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

Zhao QA, Zhou P, Goodrich R, Dmanac RT;

WPI; 2001-442253/47.

N-PSDB; AA160533.

Novel nucleic acids and polypeptides, useful for treating disorders

such as central nervous system injuries -

Example 2; SEQ ID NO 6308; 10078pp; English.

The invention relates to human nucleic acids (AA157798-AA161369) and

the encoded polypeptides (AA038642-AA042213) with nootropic,

immunosuppressant and cytostatic activity. The polynucleotides are useful

in gene therapy. A composition containing a polypeptide or polynucleotide

of the invention may be used to treat diseases of the peripheral nervous

system, such as peripheral nervous injuries, peripheral neuropathy and

localised neuropathies and central nervous system diseases, such as

Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

utilisation of the activities such as: Immune system suppression,

Activin/Inhibin activity, chemotactic/chemokinetic activity, haemostatic

and thrombolytic activity, cancer diagnosis and therapy, drug screening,

assays for receptor activity, arthritis and inflammation, leukaemias and

C.N.S disorders.

Note: The sequence data for this patent did not form part of the printed

specification.

Sequence 372 AA;

Best Local Similarity 99.7%; Pred. No. 8.7e-194;

Matches 363; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLSTLVHTYSMRYLPSVVLGAPTYVLANGVWRLSAPLPARFYQALDDRLCYGOS 60

Db 9, mlstlvhtysmrylpsvvlglcapcyvlawgywrlsalfparyqaldldrlcygys 68

QY 61 MVLFFENYTGVOILLYGDLPKKNENITLANHSTVDWIVADILAIQNALGHVRYLK 120

Db 69 mvlffenytgvoilllygdldpkknennitlanhstvdwivadilalqnalghvrylk 128

QY 121 EGLKMLPLVGCYFAQHGGIVYKSAKFNKKRKLDSYDAGIPMLVIFPESTRNPE 180

Db 129 eglkmlplvgywfyagbhgyivkysakfnekemtrklqsyvadagtpmlylviifpestrnpe 188

QY 181 QTKVLASQFAAORGAVLKHVLTPIKATHVAFDCMKVYLDAYVTVVYEGKDDGGQ 240

Db 189 qtkvlssqfaaarglavlkhvltprlkathvafdcmkkyldayvttvyyegkddggq 248

QY 241 RRESPTWTEFLCKECPKIHIDRIDKRDVPEQEHMRMLHERFEIKDKMLIEFESPD 300

Db 249 rresptwteflckeckpkihidiidridkrdvpeeqehmrmlherfeikdkmliefespd 308

QY 301 PERKKRPPGKSVNSKLSIKKTPSPMLLSGLTAGMLMTDAGRKLYVTWTYIGTLGLMY 360

Db 309 perkkrtppgksvnsksikktpsmllsgltagmlmtdagrklyvntwtlygtlglmw 368

QY 361 TIKK 364

Db 369 tika 372

RESULT 5

AA085532 standard; protein: 353 AA.

AA085532:

25-SEP-2001 (first entry)

Human secreted protein (clone id HLTH084).

Secreted protein; immunosuppressive; antiarthritic; antirheumatic;

antiproliferative; cytostatic; cardiac; vasotropic; cerebroprotective;

nootropic; neuroprotective; antibacterial; vitruclide; fungicide; human;

ophthalmological; gene therapy.

Homo sapiens.

MO200155430-A1.

02-AUG-2001.

17-JAN-2001; 2001MO-US01431.

31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 12-SEP-2000; 2000US-0231968.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Komitoulis GA, Baker KP, Birse CE, Soppet DR, Olsen HS;

Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH, Fliscella M;

Nl J, Ruben SM, Barash SC;

WPI; 2001-476220/51.

N-PSDB; AA046942.

17 isolated nucleic acid molecules encoding human secreted proteins,

used to preventing, treating or ameliorating a medical condition

Claim 11; Page 450-451; 482pp; English.

Query Match 99.4%; Score 1913; DB 22; Length 372;

CC The invention provides novel human secreted proteins and polynucleotides  
CC encoding them. The secreted proteins can be expressed by standard  
CC recombinant methodology. The secreted proteins and polynucleotides are  
CC used to prevent, treat or ameliorate a medical condition in e.g. humans,  
CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They can  
CC also be used in diagnosing a pathological condition. The antibodies to  
CC the proteins can also be used in alleviating symptoms associated with the  
CC disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme  
CC linked immunosorbent assays (ELISA). Disorders which are diagnosed or  
CC treated include autoimmune diseases e.g. rheumatoid arthritis,  
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,  
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
CC e.g. cerebral ischemia, angiodenesis, nervous system disorders e.g.  
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and  
CC ocular disorders e.g. corneal infection. The polypeptides can also be  
CC used to aid wound healing and epithelial cell proliferation, to prevent  
CC skin aging due to sunburn, to maintain organs before transplantation, for  
CC supporting cell culture of primary tissues, to regenerate tissues and in  
CC chemotaxis. The polypeptides can also be used as a food additive or  
CC preservative to increase or decrease storage capabilities. The present  
CC sequence represents a human secreted protein.

XX Sequence 353 AA;

Query Match 97.2%; Score 1871; DB 22; Length 353;  
Best Local Similarity 100.0%; Pred. No. 2.3e-189;

Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 MRYLPSVVLGTAAPYVYVWAMGWRLISAFLPARFYQALDDRLCYVQSWLFEFFENYTG 71  
DB 1 MYLLPSVVLIGTAPCYVWAWYVWLLSAFLPARFYQALDDRLCYVQSWLFEFFENYTG 60  
QY 72 VOILLTGDLPKNKENTIIYANHOSTVDWIVADILAIKRONALGHRVYVLEKGLMPLPYGC 131  
DB 61 VQILLYGDLPKKNENIIYANHOSTVDWIVADILAIKRONALGHRVYVLEKGLMPLPYGC 120  
QY 132 YFAOHGIIYVKSRAKFNKEMRNKLSYVDAGTPMYLVFPECTRYNPEQTKVLSAQAF 191  
DB 121 YFQHGIIYVKSRAKFNKEMRNKLSYVDAGTPMYLVFPECTRYNPEQTKVLSAQAF 180  
QY 192 AAOGLAVLKHVLTPIKATVAFDCMKVYLDIYDVTVVYEGKDDGGGRRSPMTTEFL 251  
DB 181 AAGRGLAVLKHVLTPIKATVAFDCMKVYLDIYDVTVVYEGKDDGGGRRSPMTTEFL 240  
QY 252 CRECPRIHIHIDRIKDDVPEDEHMRMLHEFETIKDKMLIFEYSPDERKRRPGKS 311  
DB 241 CRECPRIHIHIDRIKDDVPEDEHMRMLHEFETIKDKMLIFEYSPDERKRRPGKS 300  
QY 312 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLVTWTVYGTLLGCLWYTIKA 364  
DB 301 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLVTWTVYGTLLGCLWYTIKA 353

RESULT 6

AAB85552  
ID AAB85552 standard; protein; 353 AA.

XX AAB85552;

DT 25-SEP-2001 (first entry)

DE Human secreted protein (clone id HSL1A81).

XX Secreted protein; immunosuppressive; antiarthritic; antirheumatic;  
KW antiproliferative; cytostatic; cardiant; vasotrophic; cerebroprotective;  
KW neuroprotective; antibacterial; virucide; fungicide; human;  
XX opticalmalogical; gene therapy.

OS Homo sapiens.

XX WO200155430-A1.

XX PN

PD 02-AUG-2001.  
XX 17-JAN-2001; 2001WO-US01431.  
PF 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
XX 12-SEP-2000; 2000US-0231968.  
PA (HUMA-) HUMAN GENOME SCI INC.  
PI Rosen CA, Komatsu S, Baker KP, Birse CE, Soppet DR, Olsen HS;  
PI Moore PA, Wei P, Eder R, Duan DR, Shi Y, Choi GH, Fiscella M;  
PI N J, Ruben SM, Barash SC;  
XX MPI: 2001-476220/51.  
DR N-PSDB; AAMH6962.  
PT 17 Isolated nucleic acid molecules encoding human secreted proteins,  
PT used to preventing, treating or ameliorating a medical condition  
XX Claim 11; Page 465-466; 482pp; English.

CC The invention provides novel human secreted proteins and polynucleotides  
CC encoding them. The secreted proteins can be expressed by standard  
CC recombinant methodology. The secreted proteins and polynucleotides are  
CC used to prevent, treat or ameliorate a medical condition in e.g. humans,  
CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They can  
CC also be used in diagnosing a pathological condition. The antibodies to  
CC the proteins can also be used in alleviating symptoms associated with the  
CC disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme  
CC linked immunosorbent assays (ELISA). Disorders which are diagnosed or  
CC treated include autoimmune diseases e.g. rheumatoid arthritis,  
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,  
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
CC e.g. cerebral ischemia, angiodenesis, nervous system disorders e.g.  
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and  
CC ocular disorders e.g. corneal infection. The polypeptides can also be  
CC used to aid wound healing and epithelial cell proliferation, to prevent  
CC skin aging due to sunburn, to maintain organs before transplantation, for  
CC supporting cell culture of primary tissues, to regenerate tissues and in  
CC chemotaxis. The polypeptides can also be used as a food additive or  
CC preservative to increase or decrease storage capabilities. The present  
CC sequence represents a human secreted protein.

XX Sequence 353 AA;

Query Match 97.2%; Score 1871; DB 22; Length 353;  
Best Local Similarity 100.0%; Pred. No. 2.3e-189;

Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 MRYLPSVVLGTAAPYVYVWAMGWRLISAFLPARFYQALDDRLCYVQSWLFEFFENYTG 71  
DB 1 MYLLPSVVLIGTAPCYVWAWYVWLLSAFLPARFYQALDDRLCYVQSWLFEFFENYTG 60  
QY 72 VOILLTGDLPKNKENTIIYANHOSTVDWIVADILAIKRONALGHRVYVLEKGLMPLPYGC 131  
DB 61 VQILLYGDLPKKNENIIYANHOSTVDWIVADILAIKRONALGHRVYVLEKGLMPLPYGC 120  
QY 132 YFAOHGIIYVKSRAKFNKEMRNKLSYVDAGTPMYLVFPECTRYNPEQTKVLSAQAF 191  
DB 121 YFQHGIIYVKSRAKFNKEMRNKLSYVDAGTPMYLVFPECTRYNPEQTKVLSAQAF 180  
QY 192 AAOGLAVLKHVLTPIKATVAFDCMKVYLDIYDVTVVYEGKDDGGGRRSPMTTEFL 251  
DB 181 AAGRGLAVLKHVLTPIKATVAFDCMKVYLDIYDVTVVYEGKDDGGGRRSPMTTEFL 240  
QY 252 CRECPRIHIHIDRIKDDVPEDEHMRMLHEFETIKDKMLIFEYSPDERKRRPGKS 311  
DB 241 CRECPRIHIHIDRIKDDVPEDEHMRMLHEFETIKDKMLIFEYSPDERKRRPGKS 300  
QY 312 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLVTWTVYGTLLGCLWYTIKA 364  
DB 301 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLVTWTVYGTLLGCLWYTIKA 353

DB 301 vnsklsikktlpsmllsgltagmlmtcdagrkllyvntwlygtllgclwltika 353

RESULT 7

ID AAB93595 standard; Protein: 353 AA.

XX AAB93595;

XX 26-JUN-2001 (first entry)

XX Human protein sequence SEQ ID NO:13028.

XX Human; primer: detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

XX 11-JAN-2000; 2000JP-0116776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI: 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602

XX full-length cDNAs defined in the specification, and for the detection

XX and/or diagnosis of the abnormality of the proteins encoded by the

XX full-length cDNAs -

PS Claim 8; SEQ ID 13028; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602

XX full-length cDNAs defined in the specification. Where a primer set

XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary

XX to the complementary strand of a polynucleotide which comprises one of

XX the 5602 nucleotide sequences defined in the specification, where the

XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination

XX of an oligonucleotide comprising a sequence complementary to the

XX complementary strand of a polynucleotide which comprises a 5'-end

XX sequence and an oligonucleotide comprising a sequence complementary to a

XX polynucleotide which comprises a 3'-end sequence, where the

XX oligonucleotide comprises at least 15 nucleotides and the combination of

XX the 5'-end sequence/3'-end sequence is selected from those defined in

XX the specification. The primer sets can be used in antisense therapy and

XX in gene therapy. The primers are useful for synthesizing polynucleotides,

XX particularly full-length cDNAs. The primers are also useful for the

XX detection and/or diagnosis of the abnormality of the proteins encoded by

XX the full-length cDNAs. The primers allow obtaining of the full-length

XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and

XX AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to

XX AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632

XX represent oligonucleotides, all of which are used in the exemplification

XX of the present invention.

XX Sequence 353 AA:

Query Match 97.1%; Score 1868; DB 22; Length 353;

Best Local Similarity 99.7%; Pred. No. 4.7e-189;

Matches 352; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 12 MRYLLPSVLLGTAPTYVLAMGWKRLLSAPLPARFYQALDDRLCYQSVNLEFFENYTG 71

DB 1 mryllpsvllgtaptyvlawgvrrllsafiparfygaldrlcyqsvnmllffennytg 60

QY 72 VOILLYGDLPKNKENITLYLANHSTVDMIVADILAIRONALGHVRYLKGKMLPLXYGC 131

DB 61 vqilllygdlpknkenillylanhgstvdvialatrqnalghvrylkeglwlpilygc 120

QY 132 YFAOHGCIYVKRSKAFENKEKRNKLQSYVDGTPMYLVIFPECTRYNPEOTKVLASQAF 191

DB 121 yfaohgciyvkrsakfenekeknrlqsyvdgtpmylvifpectrynpetkvlasaqaf 180

QY 192 AAGRGGLAVLKHVLTTPRIKATHVAFDCKKNYLDALYDVTYVEGKDGQGRRESPTMEFL 251

DB 181 aagrgglavlkhlvtprrikathvafdcnknyldaldvtvyvegdkdgqgrresptmefl 240

QY 252 CKECPKIHIDRDKDQVPEDESHMRWLHERREIDKMLIEFESPDPERRRPRGKS 311

DB 241 ckecpkiihldrldkdvpeedehmrwlherreldkmliefespdpertrrprgks 300

QY 312 VNSKLSIKKTLPSMLITSGLTAGMLMTDAGRKLTVNNTWYGTLLGCLWLTIKA 364

DB 301 vnsklsikktlpsmllsgltagmlmtcdagrkllyvntwlygtllgclwltika 353

RESULT 8

AA36728

ID AAY36728 standard; Protein: 352 AA.

XX AAY36728;

XX 27-SEP-1999 (first entry)

XX Human PGI protein sequence.

XX PGI gene; biallelic marker: human; PSA; PGI-related biallelic marker:

XX cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.

XX Homo sapiens.

XX WO9932644-A2.

XX 01-JUL-1999.

XX 22-DEC-1998; 98WO-1B02133.

XX 09-SEP-1998; 98US-0099658.

XX 22-DEC-1997; 97US-0996306.

XX (GBST) GENSET.

XX Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;

XX WPI: 1999-405178/34.

XX N-PSDB: AA200872.

XX Use of a prostate cancer associated gene and biallelic markers

XX derived from it

XX Claim 7; Page 190; 385pp; English.

XX The invention relates to a mammalian PGI gene and protein, and a set of

XX PGI biallelic markers. The PGI polynucleotide and biallelic markers are

XX used in a hybridisation assay, a sequencing assay, or in an

XX allele-specific amplification assay for determining the identity of a

XX nucleotide at a PGI-related biallelic marker. The methods can be used to

XX detect and to assess the risk of developing cancer or prostate cancer.

XX Early-stage diagnosis of prostate cancer relies on prostate specific

XX antigen (PSA) dosage. However, the effectiveness of this is limited due

XX to its inability to discriminate between malignant and non-malignant

XX affections of the organ. A need exists for both a reliable diagnostic

XX procedure which would enable early-stage diagnosis, and for preventative

XX and curative treatments of the disease. The PGI gene can be used for

PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX  
DR WPI; 2001-442253/47.  
DR N-PSDB; AA158747.

PT such as central nervous system injuries -  
XX  
PS Example 4; SEQ ID NO 2736; 10078bp; English.

CC	The invention relates to human nucleic acids (AA157798-AA161369) and
CC	the encoded polypeptides (AA438642-AA442213) with neurotrophic
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful
CC	in gene therapy. A composition containing a polypeptide or polynucleotide
CC	of the invention may be used to treat diseases of the peripheral nervous
CC	system, such as peripheral nervous injuries, peripheral neuropathy and
CC	localised neuropathies and central nervous system diseases, such as
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC	utilisation of the activities such as: Immune system suppression,
CC	Activin/Inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC	assays for receptor activity, arthritis and inflammation, leukaemias and
CC	C.N.S disorders.
CC	Note: The sequence data for this patent did not form part of the printed
CC	specification.
XX	
SO	Sequence 353 AA;
Query Match	96.7%; Score 1861; DB 22; Length 353;
Best Local Similarity	99.2%; Pred. No. 2,6e-188;
Matches 350; Conservative	3; Mismatches 0; Indels 0; Gaps

[illegible]







PR	14-ANG-2000	2000US-022451.9
PR	14-ANG-2000	2000US-022521.3
PR	14-ANG-2000	2000US-022521.4
PR	14-ANG-2000	2000US-022526.7
PR	14-ANG-2000	2000US-022526.7
PR	14-ANG-2000	2000US-022526.7
PR	14-ANG-2000	2000US-022526.8
PR	14-ANG-2000	2000US-022527.0
PR	14-ANG-2000	2000US-022544.7
PR	14-ANG-2000	2000US-022545.7
PR	14-ANG-2000	2000US-022557.58
PR	14-ANG-2000	2000US-022575.59
PR	14-ANG-2000	2000US-022575.9
PR	16-ANG-2000	2000US-022627.9
PR	22-ANG-2000	2000US-022668.1
PR	22-ANG-2000	2000US-022668.681
PR	22-ANG-2000	2000US-022668.68
PR	22-ANG-2000	2000US-022717.82
PR	23-ANG-2000	2000US-022702.94
PR	30-ANG-2000	2000US-022829.24
PR	01-SEP-2000	2000US-022927.87
PR	01-SEP-2000	2000US-022934.43
PR	01-SEP-2000	2000US-022934.45
PR	01-SEP-2000	2000US-022934.45
PR	05-SEP-2000	2000US-022935.09
PR	05-SEP-2000	2000US-022935.09
PR	06-SEP-2000	2000US-022934.37
PR	06-SEP-2000	2000US-022934.37

## PR 08-5

PR 08-5  
PR 08-5

PR 08-5

PR 12-5

PR 14-5

PR 14-SEP-2000; 2000US-0232400  
PR 14-SEP-2000; 2000US-0232401  
PR 14-SEP-2000; 2000US-0233060  
PR 14-SEP-2000; 2000US-0233064  
PR 14-SEP-2000; 2000US-0233065  
PR 14-SEP-2000; 2000US-0233065  
PR 21-SEP-2000; 2000US-0234423  
PR 21-SEP-2000; 2000US-0234474  
PR 21-SEP-2000; 2000US-0234974  
PR 25-SEP-2000; 2000US-0234974  
PR 25-SEP-2000; 2000US-0234998

PR 27-5

PR 29-5

PR 29-5

PR 29-5

PR	02-OCT-2000	2000US-02436662
PR	02-OCT-2000	2000US-02370303
PR	02-OCT-2000	2000US-02370338
PR	02-OCT-2000	2000US-02373030
PR	02-OCT-2000	2000US-02373740
PR	13-OCT-2000	2000US-02393545
PR	13-OCT-2000	2000US-02393937
PR	20-OCT-2000	2000US-02439360
PR	20-OCT-2000	2000US-02418221
PR	20-OCT-2000	2000US-02418176
PR	20-OCT-2000	2000US-02417865
PR	20-OCT-2000	2000US-02417867
PR	20-OCT-2000	2000US-02418188
PR	20-OCT-2000	2000US-02418365
PR	20-OCT-2000	2000US-02418326
PR	01-NOV-2000	2000US-02446177
PR	08-NOV-2000	2000US-02464754
PR	08-NOV-2000	2000US-02464675
PR	08-NOV-2000	2000US-02466476
PR	08-NOV-2000	2000US-02466477
PR	08-NOV-2000	2000US-02466478

PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
DR WPI: 2001-457716/49.  
XX N-PSDB: AAD16620.  
XX  
PT Isolated lipid metabolism polypeptide for screening to identify  
PT antagonists and agonists that may enhance or block activities mediated  
PT by lipid metabolism proteins and also for testing and detection e.g.  
PT diagnosis -  
XX  
PS Claim 11; SEQ ID NO 47; 453bp; English.  
XX  
CC AAD16606-AAD16627 represent cDNAs corresponding to human lipid metabolism  
CC protein genes, and AAE09639-AAE09660 represent the proteins they encode.  
CC AAD16628-AAD16647 represent human lipid metabolism genomic DNAs. The  
CC genes and their corresponding proteins are useful for preventing,  
CC treating or ameliorating medical conditions such as immune disorders,  
CC inflammatory disorders (e.g., conjunctivitis, glossitis), autoimmune  
CC diseases (e.g., rheumatoid arthritis, systemic lupus erythematosus),  
CC respiratory disorders (asthma, allergy), CNS disorders (e.g., Alzheimer's  
CC disease, Parkinson's disease), haematopoietic disorders, cardiovascular  
CC disorders (e.g., atherosclerosis, coronary disease), hyperlipidemias,  
CC hypolipidaemia, lipidoses, familial hypercholesterolaemia, Gaucher's  
CC disease, Tay-Sachs's disease, diabetes mellitus, cancer and other enzyme  
CC abnormalities. They are used for the prevention of mental retardation  
CC and the enlargement of spleen and liver. They are used in gene therapy

CC and antisense therapy. The present sequence represents a human lipid  
CC metabolism protein of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX  
SQ Sequence 269 AA:  
  
Query Match 74.2%; Score 1427; DB 22; Length 269;  
Best Local Similarity 100.0%; Pred. No. 1.7e-142;  
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 96 TVDMIVADIIAIRQNALGHRVYLKGLKPLPYGCGFAOHGGIYKRSAPNEKEKRNK 155  
DB 1 tvdviwadilairqnalghrvylkeglkwlpdygcfaahgsllykrsakfnekemrnk 60  
QY 156 LQSYVDAGTPMYLVIRPEGTRYNPEQTKVLSASQAAPAOGLAVLNRHVLTPRIKATNVAP 215  
DB 61 lqsyvdagtpmylvirpegrtrynpeqtkvlssaqaafaagrlavlkvhltprikachvaf 120  
QY 216 DCKKNILDAIYDVTVYVEGKDDGGQRRESPTMEFLCKECPKHIIHDIRDKDVEPEQE 275  
DB 121 dckknlyldaiydvtlvyegkddgqqrresptmeflcckepkhihdiridkdvpeeqe 180  
QY 276 HMRRLWHERPEIKDKMLIEFYESPDEPRRRKRFPGKSVNSKLSIKTLPMSLILSGLTAGM 335  
DB 181 hmrrlwherfelkdmliefespdperrrkrfpgksvnskslsiktlpsmlisgltagm 240  
QY 336 LMTDAGRKLVTNTWYIGTLLGCLMWTIKA 364  
DB 241 lmtdagrklvntwtwyigtlgclwtelika 269  
  
RESULT 15  
AA36745  
ID AAY36745 standard; Protein; 261 AA.  
XX  
AC AAY36745;  
XX  
DT 27-SEP-1999 (first entry)  
XX  
DE PGI splice variant allele protein sequence.  
XX  
KW PGI gene; biallelic marker; human; PSA: PGI-related biallelic marker;  
KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.  
XX  
OS Homo sapiens.  
XX  
PN MO9932644-A2.  
XX  
PD 01-JUL-1999.  
XX  
PF 22-DEC-1998; 98WO-IB02133.  
XX  
PR 09-SEP-1998; 98US-0099658.  
XX  
PR 22-DEC-1997; 97US-0996306.  
XX  
PA (GEST ) GENSET.  
XX  
PI Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;  
XX  
DR WPI: 1999-405178/34.  
XX N-PSDB: AA200971.  
XX  
PT Use of a prostate cancer associated gene and biallelic markers  
PT derived from it  
XX  
PS Claim 7; Page 251; 385bp; English.  
XX  
XX The invention relates to a mammalian PGI gene and protein, and a set of  
CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
CC used in a hybridisation assay, a sequencing assay, or in an

CC allele-specific amplification assay for determining the identity of a  
CC nucleotide at a PGI-related biallelic marker. The methods can be used  
CC to detect and to assess the risk of developing cancer or prostate cancer.  
CC Early-stage diagnosis of prostate cancer relies on prostate specific  
CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
CC to its inability to discriminate between malignant and non-malignant  
CC affections of the organ. A need exists for both a reliable diagnostic  
CC procedure which would enable early-stage diagnosis, and for preventative  
CC and curative treatments of the disease. The PGI gene can be used for  
CC detection of prostate cancer, and the risk of developing it in the  
CC future, and can also be used to determine therapies for the disease.

Query Match	68.98;	Score 1325;	DB 20;	Length 261;
Post-Test Classification	73.08;	Post-Test Score 1411;		

```
Matches 261; Conservative 0; Mismatches 0; Indels 92; Gaps 1;
```

```
Oy      12 MRYLLPSVVLGTAPIYVLAWGVRLLSAFLPAPFYQALDRLCYQSMAFFIENVTG   71
         |||||
Db       1 MYLIPSVILGTAPCVLWGWRIISAFIPARFYQAIDDRLYCYSQMVIIFIEHYLG   60
```

```

0y 72 VQILLIGDLPKKNKENTIIYLANHSTVDWIVADILAIRQNALGHVRIVLKEGLKWLPLYGC 1311
    ||
Db 61 VQ----- 62

```

```

Oy 132 YFAOHGIIYKRSANEKEMHNKLOSTVDAGTEMYLVITPEEGTRINPEQTKVLSASQAF 191D
Db 63 -----mylvitpeegtrynpeqtkvlsasqaf 88

```

**Oy** 192 AAQGLAVLKHVLTPIKATHVAFDCMKNYLDATIDYVTWVEGKDGGRRRESPTWTEFL 251  
**Db** 89 aaqgslavlkhvltprikatvafdcmknyldatidvtwvyegkdggrrrespwmtefl 148			

Oy	252	CKECPKIHIDRIDKKDVPDEOEHMRRLHEREIKDKMLIEFYESPDPERRRRKREFGKS	311
Db	149	cKespKhIdrIdkKdVpeegemmrWlherteikdkmliefespdperkrfgyks	208

Qy	312	VNSKLSIKKTPSMLIS;ITAGMLMTDAGRKLYVNTWVIGTLLGCLMVTIKA	364
Db	209	VNSKLSIKKTPSMLIS;ITAGMLMTDAGRKLYVNCWVIGTLLGCLMVTIKA	261

Search completed: August 28, 2002, 11:07:34  
Job time: 478 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:08:10 ; Search time 31 Seconds  
(without alignments)  
286.804 Million cell updates/sec

Title: US-09-853-526-5  
Perfect score: 1924  
Sequence: 1 MLTSLVLTHTYSMRYLPSVY.....YVNTWIVGTLLGLMTWTFKA 364

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
Issued\_Patents\_AA:\*  
1: /cgn2\_6/prodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/prodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/prodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/prodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/prodata/2/1aa/CTUS.COMB.pep:\*  
6: /cgn2\_6/prodata/2/1aa/Backfilest1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	1924	100.0	364	2	US-08-996-306-5
2	1924	100.0	364	4	US-09-338-907-5
3	1924	100.0	364	4	US-09-218-207-5
4	1871	97.2	353	2	US-08-996-306-4
5	1871	97.2	353	4	US-09-338-907-4
6	1871	97.2	353	4	US-09-218-207-4
7	1644	85.4	315	4	US-09-338-907-134
8	1644	85.4	315	4	US-09-218-207-134
9	1568.5	81.5	354	4	US-09-338-907-74
10	1568.5	81.5	354	4	US-09-218-207-74
11	1555.5	80.8	300	4	US-09-338-907-135
12	1555.5	80.8	300	4	US-09-218-207-135
13	1498	77.9	291	4	US-09-338-907-127
14	1498	77.9	291	4	US-09-218-207-127
15	1335	68.9	261	4	US-09-338-907-128
16	1335	68.9	261	4	US-09-218-207-128
17	1203	62.5	228	4	US-09-338-907-70
18	1203	62.5	228	4	US-09-218-207-70
19	1182.5	61.5	238	4	US-09-338-907-126
20	1182.5	61.5	238	4	US-09-218-207-126
21	976	50.7	185	4	US-09-338-907-136
22	976	50.7	185	4	US-09-218-207-136
23	820	42.6	182	4	US-09-338-907-133
24	820	42.6	182	4	US-09-218-207-133
25	449	23.3	97	4	US-09-338-907-132
26	449	23.3	97	4	US-09-218-207-132
27	332	17.3	77	4	US-09-338-907-125

28	332	17.3	77	4	US-09-218-207-125	Sequence 125, App
29	330	17.2	66	4	US-09-338-907-131	Sequence 131, App
30	330	17.2	66	4	US-09-218-207-131	Sequence 131, App
31	329	17.1	68	4	US-09-338-907-130	Sequence 130, App
32	329	17.1	68	4	US-09-218-207-130	Sequence 130, App
33	329	17.1	90	4	US-09-338-907-129	Sequence 129, App
34	329	17.1	90	4	US-09-218-207-129	Sequence 129, App
35	328	17.0	374	2	US-08-454-267-2	Sequence 2, App1
36	328	17.0	374	2	US-08-941-319-2	Sequence 2, App1
37	328	17.0	374	2	US-08-941-319-6	Sequence 6, App1
38	328	17.0	374	4	US-09-035-098-2	Sequence 2, App1
39	328	17.0	374	4	US-09-035-098-6	Sequence 6, App1
40	328	17.0	374	4	US-09-215-252-5	Sequence 5, App1
41	328	17.0	374	4	US-09-215-252-17	Sequence 17, App1
42	268.5	14.0	378	4	US-09-215-252-13	Sequence 13, App1
43	240	12.5	376	4	US-08-454-267-7	Sequence 7, App1
44	239	12.4	295	2	US-08-941-319-7	Sequence 7, App1
45	239	12.4	295	2	US-08-941-319-7	Sequence 7, App1

## ALIGNMENTS

RESULT 1  
US-08-996-306-5  
; Sequence 5, Application US/08996306  
; Patent No. 5945522  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Chumakov, Ilya  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Bouguenieret, Lydie  
; TITLE OF INVENTION: Prostate cancer gene  
; NUMBER OF SEQUENCES: 68  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Knobbe, Martens, Olson & Bear  
; STREET: 501 West Broadway  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92101-3505  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Win95  
; SOFTWARE: Word  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/996,306  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israel, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: GENSET.018A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 364 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: SINGLE  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: PROTEIN  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: potential protein  
; LOCATION: 1..364  
; US-08-996-306-5

Query Match 100.0%, Score 1924, DB 2, Length 364;

```
Best Local Similarity 100.0%; Pred. No. 4.2e-200;
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLTSLVHTTYSMRYLLPSVVLGTAFTYVLAMGYWRLLSAFLPARFYQALDDRLCYCYQS 60
Db 1 MLTSLVHTTYSMRYLLPSVVLGTAFTYVLAMGYWRLLSAFLPARFYQALDDRLCYCYQS 60
QY 61 MVLFFENYTGVOILLYGDLPRKNENITYLANHSTYDVIADILAIQONLGHVRYLK 120
Db 61 MVLFFENYTGVOILLYGDLPRKNENITYLANHSTYDVIADILAIQONLGHVRYLK 120
QY 121 EGLKWLPLGYCYFAOHGIGIYKRSKAKFNEKEMRNKLOSVDAGTPMYLVIFPESTRNPE 180
Db 121 EGLKWLPLGYCYFAOHGIGIYKRSKAKFNEKEMRNKLOSVDAGTPMYLVIFPESTRNPE 180
QY 181 QTKVLSAQAFAAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDIAYDTVVEGKDDGQ 240
Db 181 QTKVLSAQAFAAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDIAYDTVVEGKDDGQ 240
QY 241 RRESPMTEFLCKECPKIHIDRIDKKDVPDEQEHMRMLHERFEIKDKMLIEFYESP 300
Db 241 RRESPMTEFLCKECPKIHIDRIDKKDVPDEQEHMRMLHERFEIKDKMLIEFYESP 300
QY 301 PERKRRPFGKSVNSKLSIKTLPMSMLISGLTAGMLMTDAGRKLIVNTWITGTLGCLMW 360
Db 301 PERKRRPFGKSVNSKLSIKTLPMSMLISGLTAGMLMTDAGRKLIVNTWITGTLGCLMW 360
QY 361 TTKA 364
Db 361 TTKA 364

RESULT 2
US-09-338-907-5
; Sequence 5, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18C11CP
; CURRENT APPLICATION NUMBER: US/09/338, 907
; EARLIER FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-03-09
; EARLIER APPLICATION NUMBER: 09/218,207
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 5
; LENGTH: 364
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-338-907-5

Query Match 100.0%; Score 1924; DB 4; Length 364;
Best Local Similarity 100.0%; Pred. No. 4.2e-200;
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Db 121 EGLKWLPLGYCYFAOHGIGIYKRSKAKFNEKEMRNKLOSVDAGTPMYLVIFPESTRNPE 180
QY 181 QTKVLSAQAFAAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDIAYDTVVEGKDDGQ 240
Db 181 QTKVLSAQAFAAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDIAYDTVVEGKDDGQ 240
QY 241 RRESPMTEFLCKECPKIHIDRIDKKDVPDEQEHMRMLHERFEIKDKMLIEFYESP 300
Db 241 RRESPMTEFLCKECPKIHIDRIDKKDVPDEQEHMRMLHERFEIKDKMLIEFYESP 300
QY 301 PERKRRPFGKSVNSKLSIKTLPMSMLISGLTAGMLMTDAGRKLIVNTWITGTLGCLMW 360
Db 301 PERKRRPFGKSVNSKLSIKTLPMSMLISGLTAGMLMTDAGRKLIVNTWITGTLGCLMW 360
QY 361 TTKA 364
Db 361 TTKA 364

RESULT 3
US-09-218-207-5
; Sequence 5, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; FILE REFERENCE: GENSET.018CP1
; CURRENT APPLICATION NUMBER: US/09/218,207
; EARLIER FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-03-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 5
; LENGTH: 364
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-218-207-5

Query Match 100.0%; Score 1924; DB 4; Length 364;
Best Local Similarity 100.0%; Pred. No. 4.2e-200;
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLTSLVHTTYSMRYLLPSVVLGTAFTYVLAMGYWRLLSAFLPARFYQALDDRLCYCYQS 60
Db 1 MLTSLVHTTYSMRYLLPSVVLGTAFTYVLAMGYWRLLSAFLPARFYQALDDRLCYCYQS 60
QY 61 MVLFFENYTGVOILLYGDLPRKNENITYLANHSTYDVIADILAIQONLGHVRYLK 120
Db 61 MVLFFENYTGVOILLYGDLPRKNENITYLANHSTYDVIADILAIQONLGHVRYLK 120
QY 121 EGLKWLPLGYCYFAOHGIGIYKRSKAKFNEKEMRNKLOSVDAGTPMYLVIFPESTRNPE 180
Db 121 EGLKWLPLGYCYFAOHGIGIYKRSKAKFNEKEMRNKLOSVDAGTPMYLVIFPESTRNPE 180
QY 181 QTKVLSAQAFAAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDIAYDTVVEGKDDGQ 240
Db 181 QTKVLSAQAFAAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDIAYDTVVEGKDDGQ 240
QY 241 RRESPMTEFLCKECPKIHIDRIDKKDVPDEQEHMRMLHERFEIKDKMLIEFYESP 300
Db 241 RRESPMTEFLCKECPKIHIDRIDKKDVPDEQEHMRMLHERFEIKDKMLIEFYESP 300
QY 301 PERKRRPFGKSVNSKLSIKTLPMSMLISGLTAGMLMTDAGRKLIVNTWITGTLGCLMW 360
Db 301 PERKRRPFGKSVNSKLSIKTLPMSMLISGLTAGMLMTDAGRKLIVNTWITGTLGCLMW 360
```

OY 361 TIRA 364  
|||  
Db 361 TIRA 364

RESULT 4  
US-08-996-306-4  
; Sequence 4, Application US/08996306  
; Patent No. 5945522  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Chumakov, Ilya  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: Prostate cancer gene  
; NUMBER OF SEQUENCES: 68  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson & Bear  
; STREET: 501 West Broadway  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92101-3505  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Win95  
; SOFTWARE: Word  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/996,306  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israelsen, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: GENSET.018A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 353 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: SINGLE  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: PROTEIN  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: potential Transmembrane helix  
; LOCATION: 1..33  
; IDENTIFICATION METHOD: Rao and Argos method  
; FEATURE:  
; NAME/KEY: potential Transmembrane helix  
; LOCATION: 4..20  
; IDENTIFICATION METHOD: Klein, Kanehisa and Delisi method  
; FEATURE:  
; NAME/KEY: potential Transmembrane helix  
; LOCATION: 4..24  
; IDENTIFICATION METHOD: Eisenberg, Schwarz, Komaromy  
; IDENTIFICATION METHOD: and Wall method  
; FEATURE:  
; NAME/KEY: potential N-myristoylation site  
; LOCATION: 12  
; IDENTIFICATION METHOD: prosite match  
; FEATURE:  
; NAME/KEY: potential Transmembrane helix  
; LOCATION: 50..70  
; IDENTIFICATION METHOD: Eisenberg, Schwarz, Komaromy and Wall method  
; FEATURE:  
; NAME/KEY: potential N-glycosylation site  
; LOCATION: 57

IDENTIFICATION METHOD: prosite match  
; FEATURE:  
; NAME/KEY: potential Transmembrane helix  
; LOCATION: 76..96  
; IDENTIFICATION METHOD: Eisenberg, Schwarz, Komaromy and Wall method  
; FEATURE:  
; NAME/KEY: potential Tyrosine kinase phosphorylation site  
; LOCATION: 78  
; IDENTIFICATION METHOD: prosite match  
; FEATURE:  
; NAME/KEY: potential Casein kinase II phosphorylation site  
; LOCATION: 84  
; IDENTIFICATION METHOD: prosite match  
; FEATURE:  
; NAME/KEY: potential Leucine zipper pattern  
; LOCATION: 94..115  
; IDENTIFICATION METHOD: prosite match  
; FEATURE:  
; NAME/KEY: potential N-myristoylation site  
; LOCATION: 119  
; IDENTIFICATION METHOD: prosite match  
; FEATURE:  
; NAME/KEY: potential Protein kinase C phosphorylation site  
; LOCATION: 133  
; IDENTIFICATION METHOD: prosite match  
; FEATURE:  
; NAME/KEY: potential Casein kinase II phosphorylation site  
; LOCATION: 147  
; IDENTIFICATION METHOD: prosite match  
; FEATURE:  
; NAME/KEY: potential Protein kinase C phosphorylation site  
; LOCATION: 194  
; IDENTIFICATION METHOD: prosite match  
; FEATURE:  
; NAME/KEY: potential Tyrosine kinase phosphorylation site  
; LOCATION: 215  
; IDENTIFICATION METHOD: prosite match  
; FEATURE:  
; NAME/KEY: potential Tyrosine sulfatation site  
; LOCATION: 221  
; IDENTIFICATION METHOD: prosite match  
; FEATURE:  
; NAME/KEY: potential cAMP- and cGMP-dependent protein kinase phosphorylation  
; LOCATION: 233  
; IDENTIFICATION METHOD: prosite match  
; FEATURE:  
; NAME/KEY: potential Casein kinase II phosphorylation site  
; LOCATION: 235  
; IDENTIFICATION METHOD: prosite match  
; FEATURE:  
; NAME/KEY: potential Protein kinase C phosphorylation site  
; LOCATION: 306  
; IDENTIFICATION METHOD: prosite match  
; FEATURE:  
; NAME/KEY: potential Transmembrane helix  
; LOCATION: 310..330  
; IDENTIFICATION METHOD: Eisenberg, Schwarz, Komaromy and Wall method  
; FEATURE:  
; NAME/KEY: potential N-myristoylation site  
; LOCATION: 319  
; IDENTIFICATION METHOD: prosite match  
; FEATURE:  
; NAME/KEY: potential N-myristoylation site  
; LOCATION: 323  
; IDENTIFICATION METHOD: prosite match  
; FEATURE:  
; NAME/KEY: potential Amidation site  
; LOCATION: 329  
; IDENTIFICATION METHOD: prosite match  
; FEATURE:  
; NAME/KEY: potential Transmembrane helix  
; LOCATION: 333..353  
; IDENTIFICATION METHOD: Eisenberg, Schwarz, Komaromy and Wall method

FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 341  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential protein kinase C phosphorylation site  
LOCATION: 350  
IDENTIFICATION METHOD: prosite match  
US-08-996-306-4

Query Match 97.2% Score 1871; DB 2; Length 355;  
Best Local Similarity 100.0%; Pred. No. 2, 2e-194;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 MRYLLPSVLLGTAPTYVLAMGVRLLSAPLPARFYQALDRLYCVYQSMVLFPEFNTG 71  
DB 1 MRYLLPSVLLGTAPTYVLAMGVRLLSAPLPARFYQALDRLYCVYQSMVLFPEFNTG 60  
QY 72 VOILLYDGLPRKKNENIYLANHSTVDMIVADIIARONALGHVRYVLEKGLMPLXGC 131  
DB 61 VOILLYDGLPRKKNENIYLANHSTVDMIVADIIARONALGHVRYVLEKGLMPLXGC 120  
QY 132 YFAOHGGIYVRSAPFKEMRNKLSQSYVDAGTDMYLVIPPEGRYRNPEDQKVLASQAF 191  
DB 121 YFAOHGGIYVRSAPFKEMRNKLSQSYVDAGTDMYLVIPPEGRYRNPEDQKVLASQAF 180  
QY 192 AAQGLAVLKHVLTPTIKATVAFDCMKKNYDDAIYDVTVVYEGKDDGQRRSPTEFL 251  
DB 181 AAQGLAVLKHVLTPTIKATVAFDCMKKNYDDAIYDVTVVYEGKDDGQRRSPTEFL 240  
QY 252 CKECPKIHIDRIDKDVPEEOEHMRMLHEFEIKDKMLIEFESDPERRRRPEGS 311  
DB 241 CKECPKIHIDRIDKDVPEEOEHMRMLHEFEIKDKMLIEFESDPERRRRPEGS 300  
QY 312 VNSKLSIKKTLPSMLIISGLAGMLMDAGRKLYVNTWYIGTLGCLMWTIKA 364  
DB 301 VNSKLSIKKTLPSMLIISGLAGMLMDAGRKLYVNTWYIGTLGCLMWTIKA 355

## RESULT 5

US-09-338-907-4  
Sequence 4, Application US/09338907  
Patent No. 6265546  
GENERAL INFORMATION:  
APPLICANT: Cohen, Daniel  
APPLICANT: Blumenfeld, Marta  
APPLICANT: Ilya, Chumakov  
APPLICANT: Bouguetoret, Lydie  
TITLE OF INVENTION: PROSTATE CANCER GENE  
FILE REFERENCE: GENSET.18CPICP  
CURRENT APPLICATION NUMBER: US/09/338.907  
CURRENT FILING DATE: 1999-06-23  
EARLIER APPLICATION NUMBER: 08/996.306  
EARLIER FILING DATE: 1997-12-22  
EARLIER APPLICATION NUMBER: 60/099.658  
EARLIER FILING DATE: 1998-09-09  
EARLIER APPLICATION NUMBER: 09/218.207  
EARLIER FILING DATE: 1998-12-22  
NUMBER OF SEQ ID NOS: 578  
SOFTWARE: Patent.pm  
SEQ ID NO 4  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 1..33  
OTHER INFORMATION: Rao and Argos identification method, potential helix  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 4..20  
OTHER INFORMATION: Klein, Kanehisa and Delist identification method, potential

FEATURE:  
NAME/KEY: HELIX  
LOCATION: 4..24  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
FEATURE:  
NAME/KEY: MYRISTATE  
LOCATION: 12..16  
OTHER INFORMATION: prosite match  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 50..70  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
FEATURE:  
NAME/KEY: CARBOHYD  
LOCATION: 57..59  
OTHER INFORMATION: prosite match  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 76..96  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 78  
OTHER INFORMATION: potential Tyrosine kinase site, prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 84  
OTHER INFORMATION: potential caseine kinase II site, prosite match  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 94..115  
OTHER INFORMATION: potential leucine zipper site, prosite match  
FEATURE:  
NAME/KEY: MYRISTATE  
LOCATION: 119..123  
OTHER INFORMATION: potential site, prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 133  
OTHER INFORMATION: potential protein kinase C, prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 147  
OTHER INFORMATION: potential caseine kinase II site, prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 194  
OTHER INFORMATION: potential protein kinase C, prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 215  
OTHER INFORMATION: potential Tyrosine kinase site, prosite match  
FEATURE:  
NAME/KEY: SULFATATION  
LOCATION: 221  
OTHER INFORMATION: prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 233  
OTHER INFORMATION: potential CAMP and cGMP dependant protein kinase site,  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 235  
OTHER INFORMATION: potential caseine kinase II site, prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 306  
OTHER INFORMATION: potential protein kinase C, prosite match  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 310..330  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
FEATURE:



```
/ NAME/KEY: MYRISTATE
/ LOCATION: 319..323
/ OTHER INFORMATION: Prositite match
/ FEATURE:
/ NAME/KEY: MYRISTATE
/ LOCATION: 323..327
/ OTHER INFORMATION: Prositite match
/ FEATURE:
/ NAME/KEY: AMIDATION
/ LOCATION: 329
/ OTHER INFORMATION: Prositite match
/ FEATURE:
/ NAME/KEY: HELIX
/ LOCATION: 333..353
/ OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F
/ FEATURE:
/ NAME/KEY: MYRISTATE
/ LOCATION: 341..345
/ OTHER INFORMATION: Prositite match
/ FEATURE:
/ NAME/KEY: PHOSPHORYLATION
/ LOCATION: 350
/ OTHER INFORMATION: potential protein kinase C, Prositite match
US-09-338-907-4
```

Query Match 97.2%; Score 1871; DB 4; Length 353;

Best Local Similarity 100.0%; Pred. No. 2.2e-194; Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 12 MRYLLPSVVLGTAPTYVYVANGVWRLLSAFLPARFYQALDDRLCYQSMVLEFFENYTG 71
DB 1 MRYLLPSVVLGTAPTYVYVANGVWRLLSAFLPARFYQALDDRLCYQSMVLEFFENYTG 60
QY 72 VOILLYGDLPRKNENIYVLANHOSYVDWIVADILAIQNALGHVRYVLEKGLKWLPLYGC 131
DB 61 VOILLYGDLPRKNENIYVLANHOSYVDWIVADILAIQNALGHVRYVLEKGLKWLPLYGC 120
QY 132 YFAOHGIVYKRSKAFENEMKRNKLOSVDAGTAPMYLVFFPGSTRNPQOTVLSAQA 191
DB 121 YFAOHGIVYKRSKAFENEMKRNKLOSVDAGTAPMYLVFFPGSTRNPQOTVLSAQA 180
QY 192 AAQRGLAVLKAVLTPRIKATHVAFPCMKNYLDIYDVTVYVYEGKDDGGRRESPTTEFL 251
DB 181 AAQRGLAVLKAVLTPRIKATHVAFPCMKNYLDIYDVTVYVYEGKDDGGRRESPTTEFL 240
QY 252 CKECPRIHIDRIDKDDVPEOEHRMRLHERFEIKDKMLIEFYSPDPERRKRPGRKS 311
DB 241 CKECPRIHIDRIDKDDVPEOEHRMRLHERFEIKDKMLIEFYSPDPERRKRPGRKS 300
QY 312 VNSKLSIKTKLPMSMLISGLTAGMLMTDAGRKLVTWTVYTGTLGCLMTTICA 364
DB 301 VNSKLSIKTKLPMSMLISGLTAGMLMTDAGRKLVTWTVYTGTLGCLMTTICA 353
```

#### RESULT 6

```
US-09-218-207-4
/ Sequence 4, Application US/09218207
/ Patent No. 6346381
/ GENERAL INFORMATION:
/ APPLICANT: Cohen, Daniel
/ APPLICANT: Blumenfeld, Marta
/ APPLICANT: Ilyu, Chumakov
/ APPLICANT: Bougueleret, Lydie
/ TITLE OF INVENTION: Prostate cancer gene
/ FILE REFERENCE: GENSET.018CPL
/ CURRENT APPLICATION NUMBER: US/09/218.207
/ EARLIER FILING DATE: 1998-12-22
/ EARLIER APPLICATION NUMBER: 08/996.306
/ EARLIER FILING DATE: 1997-12-22
/ EARLIER APPLICATION NUMBER: 60/099.658
/ EARLIER FILING DATE: 1998-09-09
/ NUMBER OF SEQ ID NOS: 578
```

```
/ SOFTWARE: Patent.pm
/ SEQ ID NO 4
/ LENGTH: 353
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: HELIX
/ LOCATION: 1..33
/ OTHER INFORMATION: Rao and Argos identification method, potential helix
/ FEATURE:
/ NAME/KEY: HELIX
/ LOCATION: 4..20
/ OTHER INFORMATION: Klein, Kanehisa and Delisi identification method, potent
/ FEATURE:
/ NAME/KEY: HELIX
/ LOCATION: 4..24
/ OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
/ FEATURE:
/ NAME/KEY: MYRISTATE
/ LOCATION: 12..16
/ OTHER INFORMATION: Prositite match
/ FEATURE:
/ NAME/KEY: HELIX
/ LOCATION: 50..70
/ OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
/ FEATURE:
/ NAME/KEY: CARBOHYD
/ LOCATION: 57..59
/ OTHER INFORMATION: Prositite match
/ FEATURE:
/ NAME/KEY: HELIX
/ LOCATION: 76..96
/ OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
/ FEATURE:
/ NAME/KEY: PHOSPHORYLATION
/ LOCATION: 78
/ OTHER INFORMATION: potential Tyrosine kinase site, Prositite match
/ FEATURE:
/ NAME/KEY: PHOSPHORYLATION
/ LOCATION: 84
/ OTHER INFORMATION: potential caseine kinase II site, Prositite match
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: 94..115
/ OTHER INFORMATION: potential Leucine zipper site, Prositite match
/ FEATURE:
/ NAME/KEY: MYRISTATE
/ LOCATION: 119..123
/ OTHER INFORMATION: potential site, Prositite match
/ FEATURE:
/ NAME/KEY: PHOSPHORYLATION
/ LOCATION: 133
/ OTHER INFORMATION: potential protein kinase C, Prositite match
/ FEATURE:
/ NAME/KEY: PHOSPHORYLATION
/ LOCATION: 147
/ OTHER INFORMATION: potential caseine kinase II site, Prositite match
/ FEATURE:
/ NAME/KEY: PHOSPHORYLATION
/ LOCATION: 194
/ OTHER INFORMATION: potential protein kinase C, Prositite match
/ FEATURE:
/ NAME/KEY: PHOSPHORYLATION
/ LOCATION: 215
/ OTHER INFORMATION: potential Tyrosine kinase site, Prositite match
/ FEATURE:
/ NAME/KEY: SULFATATION
/ LOCATION: 221
/ OTHER INFORMATION: Prositite match
/ FEATURE:
/ NAME/KEY: PHOSPHORYLATION
/ LOCATION: 233
/ OTHER INFORMATION: potential CAMP and cGMP dependant protein kinase site,
```

```

FEATURE:
NAME/KEY: PHOSPHORYLATION
LOCATION: 235
OTHER INFORMATION: potential caseine kinase II site, Prosite match
FEATURE:
NAME/KEY: PHOSPHORYLATION
LOCATION: 306
OTHER INFORMATION: potential protein kinase C, Prosite match
FEATURE:
NAME/KEY: HELIX
LOCATION: 310..330
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F
FEATURE:
NAME/KEY: MYRISTATE
LOCATION: 319..323
OTHER INFORMATION: Prosite match
FEATURE:
NAME/KEY: MYRISTATE
LOCATION: 323..327
OTHER INFORMATION: Prosite match
FEATURE:
NAME/KEY: AMIDATION
LOCATION: 329
OTHER INFORMATION: Prosite match
FEATURE:
NAME/KEY: HELIX
LOCATION: 333..353
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F
FEATURE:
NAME/KEY: MYRISTATE
LOCATION: 341..345
OTHER INFORMATION: Prosite match
FEATURE:
NAME/KEY: PHOSPHORYLATION
LOCATION: 350
OTHER INFORMATION: potential protein kinase C, Prosite match
;TS-09-218-207-4

```

```

APPLICANT: Cohen, Daniel
APPLICANT: Blumenfeld, Matia
APPLICANT: Ilya, Chumakov
APPLICANT: Bougueteler, Lydie
TITLE OF INVENTION: PROSTATE CANCER GENE
FILE REFERENCE: GENSET.18CPI.C
CURRENT APPLICATION NUMBER: US/09/338, 907
CURRENT FILING DATE: 1999-06-23
EARLIER APPLICATION NUMBER: 08/996, 306
EARLIER FILING DATE: 1997-12-22
EARLIER APPLICATION NUMBER: 06/099, 658
EARLIER FILING DATE: 1998-09-09
EARLIER APPLICATION NUMBER: 09/218, 207
EARLIER FILING DATE: 1998-12-22
NUMBER OF SEQ ID NOS: 578
SOFTWARE: Patent.pm
SEQ ID NO 134
LENGTH: 315
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: 81..83
OTHER INFORMATION: Box I
FEATURE:
NAME/KEY: SITE
LOCATION: 160..165
OTHER INFORMATION: Box II
US-09-338-907-134

```

```

; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 134
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 81..83
; OTHER INFORMATION: Box 1
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 160..165
; OTHER INFORMATION: Box 11
US-09-218-207-134
```

```

Query Match          85.4%; Score 1644; DB 4; Length 315;
Best Local Similarity 89.2%; Pred. No. 7.4e-170;
Matches 315; Conservative 0; Mismatches 0; Indels 38; Gaps 1;
```

```

QY 12 MRYLLPSVLLGTAFTYVLAAGVWRLSAFLPARFYQALDDRLCYQVSQAVLFFFEYNTG 71
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MRYLLPSVLLGTAFTYVLAAGVWRLSAFLPARFYQALDDRLCYQVSQAVLFFFEYNTG 60
QY 72 VOILLVGDLPKKNENIYYLANHOSYDWMIVADILAIQNALGHRVYLKKGKMLPLYGC 131
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 VOILLVGDLPKKNENIYYLANHOSYDWMIVADILAIQNALGHRVYLKKGKMLPLYGC 120
QY 132 YFAOHGSIYVKRSKAKNEKEMRNKLOSVDAGTPMYLVIFPEGTRVNPEDTKVLSASQAF 191
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 YFAOHGSIYVKRSKAKNEKEMRNKLOSVDAGTPMYLVIFPEGTRVNPEDTKVLSASQAF 180
QY 192 AAORGSLAVLKHVLTPRIKATHVAFDCMKNYLDAIYDVTYVVEGSKDGGRSPMTWTEFL 251
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 181 AAGRGSLAVLKHVLTPRIKATHVAFDCMKNYLDAIYDVTYVVEGSKDGGRSPMTWTEFL 202
QY 252 CKECPRKIHIDRIDKKDVEEOEHMRMWLHERFEIKDKMLIEFYESPDERRRKRPFGKS 311
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 203 CKECPRKIHIDRIDKKDVEEOEHMRMWLHERFEIKDKMLIEFYESPDERRRKRPFGKS 262
QY 312 VNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLIVNTWITIGTLGCLMWTIKA 364
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 263 VNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLIVNTWITIGTLGCLMWTIKA 315
```

```

RESULT 9
US-09-338-907-74
; Sequence 74, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CP1CP
; CURRENT APPLICATION NUMBER: US/09/338,907
; CURRENT FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218,207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 74
; LENGTH: 354
```

```

; TYPE: PRT
; ORGANISM: Mus musculus
US-09-338-907-74
```

```

Query Match          81.5%; Score 1568.5; DB 4; Length 354;
Best Local Similarity 81.4%; Pred. No. 1.4e-161;
Matches 288; Conservative 39; Mismatches 26; Indels 1; Gaps 1;
```

```

QY 12 MRYLLPSVLLGTAFTYVLAAGVWRLSAFLPARFYQALDDRLCYQVSQAVLFFFEYNTG 71
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MRYLLPSVLLGTAFTYVLAAGVWRLSAFLPARFYQALDDRLCYQVSQAVLFFFEYNTG 60
QY 72 VOILLVGDLPKKNENIYYLANHOSYDWMIVADILAIQNALGHRVYLKKGKMLPLYGC 131
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 VOILLVGDLPKKNENIYYLANHOSYDWMIVADILAIQNALGHRVYLKKGKMLPLYGC 120
QY 132 YFAOHGSIYVKRSKAKNEKEMRNKLOSVDAGTPMYLVIFPEGTRVNPEDTKVLSASQAF 191
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 YFAOHGSIYVKRSKAKNEKEMRNKLOSVDAGTPMYLVIFPEGTRVNPEDTKVLSASQAF 180
QY 192 AAORGSLAVLKHVLTPRIKATHVAFDCMKNYLDAIYDVTYVVEGSKDG-GGRSPMTWTEFL 250
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 181 AAGRGSLAVLKHVLTPRIKATHVAFDSKSHLDAIYDVTYVVEGSKDGSKSPMTWTEFL 240
QY 251 CKECPRKIHIDRIDKKDVEEOEHMRMWLHERFEIKDKMLIEFYESPDERRRKRPFGK 310
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 241 CKOCPRKIHIDRIDKNEVPEOEHMRMWLHERFEIKDKMLIEFYDSDPERRRKRPFGK 300
QY 311 SVNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLIVNTWITIGTLGCLMWTIKA 364
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 301 SVNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLIVNTWITIGTLGCLMWTIKA 354
```

```

RESULT 10
US-09-218-207-74
; Sequence 74, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; FILE REFERENCE: GENSET.018CP1
; CURRENT APPLICATION NUMBER: US/09/218,207
; CURRENT FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 74
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-218-207-74
```

```

Query Match          81.5%; Score 1568.5; DB 4; Length 354;
Best Local Similarity 81.4%; Pred. No. 1.4e-161;
Matches 288; Conservative 39; Mismatches 26; Indels 1; Gaps 1;
```

```

QY 12 MRYLLPSVLLGTAFTYVLAAGVWRLSAFLPARFYQALDDRLCYQVSQAVLFFFEYNTG 71
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MRYLLPSVLLGTAFTYVLAAGVWRLSAFLPARFYQALDDRLCYQVSQAVLFFFEYNTG 60
QY 72 VOILLVGDLPKKNENIYYLANHOSYDWMIVADILAIQNALGHRVYLKKGKMLPLYGC 131
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 VOILLVGDLPKKNENIYYLANHOSYDWMIVADILAIQNALGHRVYLKKGKMLPLYGC 120
QY 132 YFAOHGSIYVKRSKAKNEKEMRNKLOSVDAGTPMYLVIFPEGTRVNPEDTKVLSASQAF 191
```

```
|||||
Db 121 YFAOHGSIYKRSKAFNDKEMRSKLSQSYVNAAGTPTMYLVIFPEEGTRYNATYTKILSASQAF 180
Qy 192 AAORGGLAVLHVLPRIKATHVAFDCKKNYLDAIYDVTYVYEGKDDG-GGQRESPTMEF 250
Db 181 AAORGGLAVLHVLPRIKATHVAFDCKSHLDALYDVTYVYEGKESGKYSNPSPMTEF 240
Qy 251 LCKECPRIHIHIDRIKDVPEEGEHMRRLHEREFIKDKMLIEFESPPDERRRKRPFGK 310
Db 241 LCKGCPRLHIHFDRIIDNEVPEEGEHMKMLHEREFIKDRLLIEFYDPSDERINKRPFPGK 300
Qy 311 SVNSKLSIKTKLPMSMLILSGLTAGMLMTDAGRKLVTWNTWYIGTLGCLWVTIKA 364
Db 301 SVHSRLSVKTKLPMSVLILGSLTAVLMTESGRKLYMGTMWYIGTLGCLWFEVYIKA 354

RESULT 11
US-09-338-907-135
; Sequence 135, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marla
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CP1CP
; CURRENT APPLICATION NUMBER: US/09/338,907
; EARLIER FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218,207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 135
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 81..83
; OTHER INFORMATION: Box I
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 160..165
; OTHER INFORMATION: Box II
US-09-338-907-135

Query Match 80.8%; Score 1555.5; DB 4; Length 300;
Best Local Similarity 85.0%; Pred. No.2.7e-160;
Matches 300; Conservative 0; Mismatches 0; Indels 53; Gaps 1;
```

```
|||||
Db 188 CKECPKIHIDRIKDKVDPEEGEHMRRLHEREFIKDKMLIEFESPPDERRRKRPFGKS 247
Qy 312 VNSKLSIKTKLPMSMLILSGLTAGMLMTDAGRKLVTWNTWYIGTLGCLWVTIKA 364
Db 248 VNSKLSIKTKLPMSMLILSGLTAGMLMTDAGRKLVTWNTWYIGTLGCLWVTIKA 300

RESULT 12
US-09-218-207-135
; Sequence 135, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marla
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; FILE REFERENCE: GENSET.018CP1
; CURRENT APPLICATION NUMBER: US/09/218,207
; EARLIER FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 135
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 81..83
; OTHER INFORMATION: Box I
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 160..165
; OTHER INFORMATION: Box II
US-09-218-207-135

Query Match 80.8%; Score 1555.5; DB 4; Length 300;
Best Local Similarity 85.0%; Pred. No.2.7e-160;
Matches 300; Conservative 0; Mismatches 0; Indels 53; Gaps 1;
```

```
; Sequence 127, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CPLCP
; CURRENT APPLICATION NUMBER: US/09/338,907
; CURRENT FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218,207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 127
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 98..103
; OTHER INFORMATION: Box II
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 149..157
; OTHER INFORMATION: Box III
US-09-338-907-127
```

```
Query Match      77.9%; Score 1498; DB 4; Length 291;
Best Local Similarity 82.4%; Pred. No. 4,4e-154;
Matches 291; Conservative 0; Mismatches 0; Indels 62; Gaps 1;

QY 12 MRYLLPSVLLGTAPTYVLANGVWRLLSAFLPARFYQALDRLCYVOSWLPFFENYTG 71
    |||||||
DB 1 MRYLLPSVLLGTAPTYVLANGVWRLLSAFLPARFYQALDRLCYVOSWLPFFENYTG 60
    |||||||
QY 72 VOILLYGDLPKKNENIYYLANHOSTVDIVADIIAIRONALGHRYVYLKEGLKMLPLYGC 131
    |
DB 61 V----- 61
QY 132 YFAOHGIIYVRSKAKFENEKEMRNKLOSVDAGTPTMYLVIFPEGTRYNPEOTKVLASQAF 191
    |||||||
DB 62 ---QHGGIYVRSKAKFENEKEMRNKLOSVDAGTPTMYLVIFPEGTRYNPEOTKVLASQAF 118
    |||||||
QY 192 AAGRGALVAKHVLTPRIKATHVAFDCMKNYLDATYDTVYVEGDDGGQRRRESPTMTFFL 251
    |||||||
DB 119 AAGRGALVAKHVLTPRIKATHVAFDCMKNYLDATYDTVYVEGDDGGQRRRESPTMTFFL 178
    |||||||
QY 252 CKECPKTHIHIDRIDKRDVPEOEHEMRMWLHEREFEIKDKMLIEFYESPDPERRRRPFQGS 311
    |||||||
DB 179 CKECPKTHIHIDRIDKRDVPEOEHEMRMWLHEREFEIKDKMLIEFYESPDPERRRRPFQGS 238
    |||||||
QY 312 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLYVNTWYIGTLGCLMTVITKA 364
    |||||||
DB 239 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLYVNTWYIGTLGCLMTVITKA 291
    |||||||
```

```
RESULT 14
US-09-218-207-127
; Sequence 127, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
```

```
; FILE REFERENCE: GENSET.018CPI
; CURRENT APPLICATION NUMBER: US/09/218,207
; CURRENT FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 127
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 98..103
; OTHER INFORMATION: Box II
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 149..157
; OTHER INFORMATION: Box III
US-09-218-207-127
```

```
Query Match      77.9%; Score 1498; DB 4; Length 291;
Best Local Similarity 82.4%; Pred. No. 4,4e-154;
Matches 291; Conservative 0; Mismatches 0; Indels 62; Gaps 1;

QY 12 MRYLLPSVLLGTAPTYVLANGVWRLLSAFLPARFYQALDRLCYVOSWLPFFENYTG 71
    |||||||
DB 1 MRYLLPSVLLGTAPTYVLANGVWRLLSAFLPARFYQALDRLCYVOSWLPFFENYTG 60
    |||||||
QY 72 VOILLYGDLPKKNENIYYLANHOSTVDIVADIIAIRONALGHRYVYLKEGLKMLPLYGC 131
    |
DB 61 V----- 61
QY 132 YFAOHGIIYVRSKAKFENEKEMRNKLOSVDAGTPTMYLVIFPEGTRYNPEOTKVLASQAF 191
    |||||||
DB 62 ---QHGGIYVRSKAKFENEKEMRNKLOSVDAGTPTMYLVIFPEGTRYNPEOTKVLASQAF 118
    |||||||
QY 192 AAGRGALVAKHVLTPRIKATHVAFDCMKNYLDATYDTVYVEGDDGGQRRRESPTMTFFL 251
    |||||||
DB 119 AAGRGALVAKHVLTPRIKATHVAFDCMKNYLDATYDTVYVEGDDGGQRRRESPTMTFFL 178
    |||||||
QY 252 CKECPKTHIHIDRIDKRDVPEOEHEMRMWLHEREFEIKDKMLIEFYESPDPERRRRPFQGS 311
    |||||||
DB 179 CKECPKTHIHIDRIDKRDVPEOEHEMRMWLHEREFEIKDKMLIEFYESPDPERRRRPFQGS 238
    |||||||
QY 312 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLYVNTWYIGTLGCLMTVITKA 364
    |||||||
DB 239 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLYVNTWYIGTLGCLMTVITKA 291
    |||||||
```

```
RESULT 15
US-09-338-907-128
; Sequence 128, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CPLCP
; CURRENT APPLICATION NUMBER: US/09/338,907
; CURRENT FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218,207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
```

```
; SOFTWARE: Patent.pm
; SEQ ID NO 128
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 68..73
; OTHER INFORMATION: Box II
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 119..127
; OTHER INFORMATION: Box III
; US-09-338-907-128
```

```
Query Match          68.9%; Score 1325; DB 4; length 261:
Best Local Similarity 73.9%; Pred. No. 2,1e-135;
Matches 261; Conservative 0; Mismatches 0; Indels 92; Gaps 1;
```

```
QY 12 MRYLLPSVVLGTAPTYVLANGVNRLLSAFLPARFYQALDDRLCYQSMVLEFFENYTG 71
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MRYLLPSVVLGTAPTYVLANGVNRLLSAFLPARFYQALDDRLCYQSMVLEFFENYTG 60

QY 72 VQILLYGDLPRNKENIYLANHGSTVDWIVADILAIQNALGHVARYLKEGIKMLPLYGC 131
    ||
Db 61 VQ----- 62

QY 132 YFAQHGIYVRSKAFNENKEMRNKLQSYVDAGTPIWYVIFPEGTRYNPEOTKVLASQAF 191
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 63 -----MYLVIFPEGTRYNPEOTKVLASQAF 88

QY 192 AAQGLAVLKHLVTPRIKATHVAFDCMKKNYLDAYDVYVEGKDDGGQRRSEPTMEFL 251
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 89 AAQGLAVLKHLVTPRIKATHVAFDCMKKNYLDAYDVYVEGKDDGGQRRSEPTMEFL 148

QY 252 CKECPKIHIDRIDDKDVPEEQEHMRWLHERFEIKDKMLIEFYESPDPERRKRFPGKS 311
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 149 CKECPKIHIDRIDDKDVPEEQEHMRWLHERFEIKDKMLIEFYESPDPERRKRFPGKS 208

QY 312 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTNTWYIGTLGCLMWTIKA 364
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 209 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTNTWYIGTLGCLMWTIKA 261
```

Search completed: August 28, 2002, 11:08:11  
Job time: 290 sec



Db 61 MVLFFENYTGVOILLYGDLPKKNENIITYLANHOSTVDWIVADILAIROMALGHVRYLK 120  
Qy 121 EGLKWLPLGYCYFAOHGCIYVYKRSKAFNEKEMRNKLOSVDAGTAPMYLVIFPESTRYNPE 180  
Db 121 EGLKWLPLGYCYFAOHGCIYVYKRSKAFNEKEMRNKLOSVDAGTAPMYLVIFPESTRYNPE 180  
Qy 181 QTKVLSASQAFAAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVYVECKDDGGQ 240  
Db 181 QTKVLSASQAFAAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVYVECKDDGGQ 240  
Qy 241 RRESPMTEFLCKECPRIHIHIDRIKDDVPEEOEHMRMLHEFEIKDKMLIEFYESP 300  
Db 241 RRESPMTEFLCKECPRIHIHIDRIKDDVPEEOEHMRMLHEFEIKDKMLIEFYESP 300  
Qy 301 PERKRRPGKSVNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLIVMTWITIGTLGCLMW 360  
Db 301 PERKRRPGKSVNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLIVMTWITIGTLGCLMW 360  
Qy 361 TIKK 364  
Db 361 TIKK 364

## RESULT 2

US-09-817-910-7  
; Sequence 7, Application US/09817910  
; GENERAL INFORMATION:  
; APPLICANT: Meyers, Rachel A.  
; APPLICANT: Macbeth, Kyle J.  
; APPLICANT: Williamson, Mark  
; APPLICANT: Rudolph-Owen, Laura A.  
; APPLICANT: Tsai, Fong-Ying  
; TITLE OF INVENTION: 46743 and 27417 NOVEL HUMAN  
; TITLE OF INVENTION: ACYLTRANSFEROSES AND USES THEREFOR  
; FILE REFERENCE: 10448-032001  
; CURRENT APPLICATION NUMBER: US/09/817, 910  
; CURRENT FILING DATE: 2001-03-26  
; PRIOR APPLICATION NUMBER: US 60/192,092  
; PRIOR FILING DATE: 2000-03-24  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 364  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-817-910-7

Query Match 100.0%; Score 1924; DB 22; Length 364;

Best Local Similarity 100.0%; Pred. No. 4e-189;

Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVLFFENYTGVOILLYGDLPKKNENIITYLANHOSTVDWIVADILAIROMALGHVRYLK 60  
Db 1 MVLFFENYTGVOILLYGDLPKKNENIITYLANHOSTVDWIVADILAIROMALGHVRYLK 60  
Qy 61 MVLFFENYTGVOILLYGDLPKKNENIITYLANHOSTVDWIVADILAIROMALGHVRYLK 120  
Db 61 MVLFFENYTGVOILLYGDLPKKNENIITYLANHOSTVDWIVADILAIROMALGHVRYLK 120  
Qy 121 EGLKWLPLGYCYFAOHGCIYVYKRSKAFNEKEMRNKLOSVDAGTAPMYLVIFPESTRYNPE 180  
Db 121 EGLKWLPLGYCYFAOHGCIYVYKRSKAFNEKEMRNKLOSVDAGTAPMYLVIFPESTRYNPE 180  
Qy 181 QTKVLSASQAFAAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVYVECKDDGGQ 240  
Db 181 QTKVLSASQAFAAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVYVECKDDGGQ 240  
Qy 241 RRESPMTEFLCKECPRIHIHIDRIKDDVPEEOEHMRMLHEFEIKDKMLIEFYESP 300  
Db 241 RRESPMTEFLCKECPRIHIHIDRIKDDVPEEOEHMRMLHEFEIKDKMLIEFYESP 300  
Qy 301 PERKRRPGKSVNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLIVMTWITIGTLGCLMW 360  
Db 301 PERKRRPGKSVNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLIVMTWITIGTLGCLMW 360

Db 301 PERKRRPGKSVNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLIVMTWITIGTLGCLMW 360  
Qy 361 TIKK 364  
Db 361 TIKK 364

## RESULT 3

US-09-853-526-5  
; Sequence 5, Application US/09853526  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Ilya, Chumakov  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: PROSTATE CANCER GENE  
; FILE REFERENCE: GENSET.18CPICP  
; CURRENT APPLICATION NUMBER: US/09/853, 526  
; CURRENT FILING DATE: 2001-05-11  
; PRIOR APPLICATION NUMBER: 09/338,907  
; PRIOR FILING DATE: 1999-06-23  
; PRIOR APPLICATION NUMBER: 08/996,306  
; PRIOR FILING DATE: 1997-12-22  
; PRIOR APPLICATION NUMBER: 60/099,658  
; PRIOR FILING DATE: 1998-09-09  
; PRIOR APPLICATION NUMBER: 09/218,207  
; PRIOR FILING DATE: 1998-12-22  
; NUMBER OF SEQ ID NOS: 578  
; SOFTWARE: Patent.pm  
; SEQ ID NO 5  
; LENGTH: 364  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-853-526-5

Query Match 100.0%; Score 1924; DB 22; Length 364;

Best Local Similarity 100.0%; Pred. No. 4e-189;

Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVLFFENYTGVOILLYGDLPKKNENIITYLANHOSTVDWIVADILAIROMALGHVRYLK 60  
Db 1 MVLFFENYTGVOILLYGDLPKKNENIITYLANHOSTVDWIVADILAIROMALGHVRYLK 60  
Qy 61 MVLFFENYTGVOILLYGDLPKKNENIITYLANHOSTVDWIVADILAIROMALGHVRYLK 120  
Db 61 MVLFFENYTGVOILLYGDLPKKNENIITYLANHOSTVDWIVADILAIROMALGHVRYLK 120  
Qy 121 EGLKWLPLGYCYFAOHGCIYVYKRSKAFNEKEMRNKLOSVDAGTAPMYLVIFPESTRYNPE 180  
Db 121 EGLKWLPLGYCYFAOHGCIYVYKRSKAFNEKEMRNKLOSVDAGTAPMYLVIFPESTRYNPE 180  
Qy 181 QTKVLSASQAFAAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVYVECKDDGGQ 240  
Db 181 QTKVLSASQAFAAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVYVECKDDGGQ 240  
Qy 241 RRESPMTEFLCKECPRIHIHIDRIKDDVPEEOEHMRMLHEFEIKDKMLIEFYESP 300  
Db 241 RRESPMTEFLCKECPRIHIHIDRIKDDVPEEOEHMRMLHEFEIKDKMLIEFYESP 300  
Qy 301 PERKRRPGKSVNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLIVMTWITIGTLGCLMW 360  
Db 301 PERKRRPGKSVNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLIVMTWITIGTLGCLMW 360  
Qy 361 TIKK 364  
Db 361 TIKK 364

## RESULT 4

US-09-901-484A-5  
; Sequence 5, Application US/09901484A



```

; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate Cancer Gene
; FILE REFERENCE: GEN-T11XC3D2
; CURRENT APPLICATION NUMBER: US/09/901,484A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 08/996,306
; PRIOR FILING DATE: 1997-12-22
; PRIOR APPLICATION NUMBER: US 60/099,658
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: US 09/218,207
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: US 09/338,907
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: US 09/853,526
; PRIOR FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 5
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-901-484A-5

```

```

Query Match      100.0%; Score 1924; DB 23; Length 364;
Best Local Similarity 100.0%; Pred. No. 4e-189;
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MLISLVHTYSMRYLLPSVVLGTAPTYVLAMGVWRLLSAFLPARFYQALDRLCYVOS 60
DB 1 MLISLVHTYSMRYLLPSVVLGTAPTYVLAMGVWRLLSAFLPARFYQALDRLCYVOS 60
QY 61 MYLFFENTYGVQIILLYGDLPRKNENIYLANHOSVDMIVADILAIRONALGHVRYLX 120
DB 61 MYLFFENTYGVQIILLYGDLPRKNENIYLANHOSVDMIVADILAIRONALGHVRYLX 120
QY 121 EGIKWLPLYGCVFAOHGIIYVRSKAFNEKEMNKLOSVDAGTPMYLVIFPEGTRYNPE 180
DB 121 EGIKWLPLYGCVFAOHGIIYVRSKAFNEKEMNKLOSVDAGTPMYLVIFPEGTRYNPE 180
QY 181 QFVLASASQAFQAORGLAVLKHVLPRIKATHVAFDCKMNYDAIYDVYVYEGKDDGGQ 240
DB 181 QFVLASASQAFQAORGLAVLKHVLPRIKATHVAFDCKMNYDAIYDVYVYEGKDDGGQ 240
QY 241 RRESPTMTEFLCKECKPKIHIDRIDKDVPEBOEHMRRLHREFIKDKMLIEFYESP 300
DB 241 RRESPTMTEFLCKECKPKIHIDRIDKDVPEBOEHMRRLHREFIKDKMLIEFYESP 300
QY 301 PERRKRPFGKSVNSKLSIKKTLPMSLILSGITAGMLMTDAGRKLVTYNTWYIGTLGCLMV 360
DB 301 PERRKRPFGKSVNSKLSIKKTLPMSLILSGITAGMLMTDAGRKLVTYNTWYIGTLGCLMV 360
QY 361 TTKA 364
DB 361 TTKA 364

```

```

RESULT 5
US-60-099-658-5
; Sequence 5, Application US/60099658
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Chumakov, Ilya
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; NUMBER OF SEQUENCES: 99
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear

```

```

; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/099,658
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelson, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: GENSET.018APR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 364 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: potential protein
; LOCATION: 1..364
US-60-099-658-5

```

```

Query Match      100.0%; Score 1924; DB 26; Length 364;
Best Local Similarity 100.0%; Pred. No. 4e-189;
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MLISLVHTYSMRYLLPSVVLGTAPTYVLAMGVWRLLSAFLPARFYQALDRLCYVOS 60
DB 1 MLISLVHTYSMRYLLPSVVLGTAPTYVLAMGVWRLLSAFLPARFYQALDRLCYVOS 60
QY 61 MYLFFENTYGVQIILLYGDLPRKNENIYLANHOSVDMIVADILAIRONALGHVRYLX 120
DB 61 MYLFFENTYGVQIILLYGDLPRKNENIYLANHOSVDMIVADILAIRONALGHVRYLX 120
QY 121 EGIKWLPLYGCVFAOHGIIYVRSKAFNEKEMNKLOSVDAGTPMYLVIFPEGTRYNPE 180
DB 121 EGIKWLPLYGCVFAOHGIIYVRSKAFNEKEMNKLOSVDAGTPMYLVIFPEGTRYNPE 180
QY 181 QFVLASASQAFQAORGLAVLKHVLPRIKATHVAFDCKMNYDAIYDVYVYEGKDDGGQ 240
DB 181 QFVLASASQAFQAORGLAVLKHVLPRIKATHVAFDCKMNYDAIYDVYVYEGKDDGGQ 240
QY 241 RRESPTMTEFLCKECKPKIHIDRIDKDVPEBOEHMRRLHREFIKDKMLIEFYESP 300
DB 241 RRESPTMTEFLCKECKPKIHIDRIDKDVPEBOEHMRRLHREFIKDKMLIEFYESP 300
QY 301 PERRKRPFGKSVNSKLSIKKTLPMSLILSGITAGMLMTDAGRKLVTYNTWYIGTLGCLMV 360
DB 301 PERRKRPFGKSVNSKLSIKKTLPMSLILSGITAGMLMTDAGRKLVTYNTWYIGTLGCLMV 360
QY 361 TTKA 364
DB 361 TTKA 364

```

```

RESULT 6
US-09-488-725A-6308
; Sequence 6308, Application US/09488725A
; GENERAL INFORMATION:

```

```
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides
; FILE REFERENCE: 784FLPCT
; CURRENT APPLICATION NUMBER: US/09/488, 725A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US/09/488, 725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552, 317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US09/598, 042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: US09/620, 312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: US09/653, 450
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US09/662, 191
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US09/693, 036
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US09/727, 344
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7144
; SOFTWARE: PL_FL_genes_b Versions 1.0
; SEQ ID NO 6308
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-488-725A-6308

Query Match      99.4%; Score 1913; DB 18; Length 372;
Best Local Similarity 99.7%; Pred. No. 5,7e-188;
Matches 363; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLSLVLTHTYSMRVLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQSVYQS 60
DB 9 MLSLVLTHTYSMRVLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQSVYQS 68
QY 61 MVLFFENYTGVOILLYGLDLPKKNENITYLANHOSYVDWIVADILAIROMLGHVRVYLK 120
DB 69 MVLFFENYTGVOILLYGLDLPKKNENITYLANHOSYVDWIVADILAIROMLGHVRVYLK 128
QY 121 EGLWVLPYGCYFAOHGIIYVRSKAKENEKMRNKLOSVDAGTPMYLVIFPEGTRYNPE 180
DB 129 EGLWVLPYGCYFAOHGIIYVRSKAKENEKMRNKLOSVDAGTPMYLVIFPEGTRYNPE 188
QY 181 QTKVLSAQAFAOAGLAVLKHVLTLPRIKATHVAFDCMKNYLDIYDVTVYVEGKDDGGO 240
DB 189 QTKVLSAQAFAOAGLAVLKHVLTLPRIKATHVAFDCMKNYLDIYDVTVYVEGKDDGGO 248
QY 241 RRESPMTTEFLCKECPKIHIDRIDKKDVEPEQEHMRMWLHERFEIKDKMLIEFYESPD 300
DB 249 RRESPMTTEFLCKECPKIHIDRIDKKDVEPEQEHMRMWLHERFEIKDKMLIEFYESPD 308
QY 301 PERKKRPPGKSVNSKLSIKTLPMSLLISGLTAGMLMTDAGRKLIVNTWIIYGTLLGLCLWV 360
DB 309 PERKKRPPGKSVNSKLSIKTLPMSLLISGLTAGMLMTDAGRKLIVNTWIIYGTLLGLCLWV 368
QY 361 TIRKA 364
DB 369 TIRKA 372

RESULT 7
PCT-US01-01431-59
; Sequence 59, Application PC/TUS0101431
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: 17 human secreted proteins
; FILE REFERENCE: PS723PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01431
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179, 065
```

```
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180, 628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/231, 968
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 59
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-01431-59

Query Match      97.2%; Score 1871; DB 1; Length 353;
Best Local Similarity 100.0%; Pred. No. 1.1e-183;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 MRYLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQSVYQSMVLEFFENYTG 71
DB 1 MRYLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQSVYQSMVLEFFENYTG 60
QY 72 VOILLYGLDLPKKNENITYLANHOSYVDWIVADILAIROMLGHVRVYLKKGKMLPLYGC 121
DB 61 VOILLYGLDLPKKNENITYLANHOSYVDWIVADILAIROMLGHVRVYLKKGKMLPLYGC 120
QY 132 YFAOHGIIYVRSKAKENEKMRNKLOSVDAGTPMYLVIFPEGTRYNPEQTKVLSAQAFA 191
DB 121 YFAOHGIIYVRSKAKENEKMRNKLOSVDAGTPMYLVIFPEGTRYNPEQTKVLSAQAFA 180
QY 192 AAORGGLAVLKHVLTLPRIKATHVAFDCMKNYLDIYDVTVYVEGKDDGGO RRESPMTTEFL 251
DB 181 AAORGGLAVLKHVLTLPRIKATHVAFDCMKNYLDIYDVTVYVEGKDDGGO RRESPMTTEFL 240
QY 252 CKECRPHIHIDRIDKKDVEPEQEHMRMWLHERFEIKDKMLIEFYESPDERRRRPPGKS 311
DB 241 CKECRPHIHIDRIDKKDVEPEQEHMRMWLHERFEIKDKMLIEFYESPDERRRRPPGKS 300
QY 312 VNSKLSIKTLPMSLLISGLTAGMLMTDAGRKLIVNTWIIYGTLLGLCLWV TIRKA 364
DB 301 VNSKLSIKTLPMSLLISGLTAGMLMTDAGRKLIVNTWIIYGTLLGLCLWV TIRKA 353

RESULT 8
PCT-US01-01431-79
; Sequence 79, Application PC/TUS0101431
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: 17 human secreted proteins
; FILE REFERENCE: PS723PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01431
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179, 065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180, 628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/231, 968
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 79
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-01431-79

Query Match      97.2%; Score 1871; DB 1; Length 353;
Best Local Similarity 100.0%; Pred. No. 1.1e-183;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 MRYLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQSVYQSMVLEFFENYTG 71
DB 1 MRYLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQSVYQSMVLEFFENYTG 71
```

```
Db      1  MRLLPSVVLGTAPTYVVLAMGWRLLSAFLPARFYQALDDRLCYCYQSQWVLEFFENYTG 60
QY      72  VQILLYGDLPKKNENIYYLANHSTVDWIADILAIROMALGHVRYLKEGLKMLPLYGC 131
Db      61  VQILLYGDLPKKNENIYYLANHSTVDWIADILAIROMALGHVRYLKEGLKMLPLYGC 120
QY      132 YFAOHGGIYKRSKAKNEKEMRNKLOSVDAGTPMYLVIFPESTRNPEQTKYLSASQAF 191
Db      121 YFAOHGGIYKRSKAKNEKEMRNKLOSVDAGTPMYLVIFPESTRNPEQTKYLSASQAF 180
QY      192 AAOAGLAVLKHLVLPRIKATHVAFDCMKNYLDIYDVTYVVEGKDDGGQRESPTMTTEFL 251
Db      181 AAOAGLAVLKHLVLPRIKATHVAFDCMKNYLDIYDVTYVVEGKDDGGQRESPTMTTEFL 240
QY      252 CKCEPKIHIDRIDKKDVEEEOEHMRWMLHERFEIKDKMLIEFYSPDERRRKRPFGKS 311
Db      241 CKCEPKIHIDRIDKKDVEEEOEHMRWMLHERFEIKDKMLIEFYSPDERRRKRPFGKS 300
QY      312 VNSKLSIKTLPBMLILSGLTAGMLMTDAGRKLVTWNTWYIGTLGCLMTWYTIKA 364
Db      301 VNSKLSIKTLPBMLILSGLTAGMLMTDAGRKLVTWNTWYIGTLGCLMTWYTIKA 353
```

```
RESULT 9
PCT-US01-11988-1474
; Sequence 1474, Application PC/TUS0111988
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Human Genome Sciences, Inc.
```

```
; TITLE OF INVENTION: Albumin Fusion Proteins
```

```
; FILE REFERENCE: PF546PCT
```

```
; CURRENT APPLICATION NUMBER: PCT/US01/11988
```

```
; CURRENT FILING DATE: 2001-01-12
```

```
; PRIOR APPLICATION NUMBER: 60/229, 358
```

```
; PRIOR FILING DATE: 2000-04-12
```

```
; PRIOR APPLICATION NUMBER: 60/256, 931
```

```
; PRIOR FILING DATE: 2000-12-21
```

```
; PRIOR APPLICATION NUMBER: 60/199, 384
```

```
; PRIOR FILING DATE: 2000-04-25
```

```
; NUMBER OF SEQ ID NOS: 2267
```

```
; SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 1474
```

```
; LENGTH: 353
```

```
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
```

```
PCT-US01-11988-1474
```

```
Query Match          97.2%; Score 1871; DB 1; Length 353;
Best Local Similarity 100.0%; Pred. No. 1.1e-183;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      12  MRLLPSVVLGTAPTYVVLAMGWRLLSAFLPARFYQALDDRLCYCYQSQWVLEFFENYTG 71
Db      1  MRLLPSVVLGTAPTYVVLAMGWRLLSAFLPARFYQALDDRLCYCYQSQWVLEFFENYTG 60
QY      72  VQILLYGDLPKKNENIYYLANHSTVDWIADILAIROMALGHVRYLKEGLKMLPLYGC 131
Db      61  VQILLYGDLPKKNENIYYLANHSTVDWIADILAIROMALGHVRYLKEGLKMLPLYGC 120
QY      132 YFAOHGGIYKRSKAKNEKEMRNKLOSVDAGTPMYLVIFPESTRNPEQTKYLSASQAF 191
Db      121 YFAOHGGIYKRSKAKNEKEMRNKLOSVDAGTPMYLVIFPESTRNPEQTKYLSASQAF 180
QY      192 AAOAGLAVLKHLVLPRIKATHVAFDCMKNYLDIYDVTYVVEGKDDGGQRESPTMTTEFL 251
Db      181 AAOAGLAVLKHLVLPRIKATHVAFDCMKNYLDIYDVTYVVEGKDDGGQRESPTMTTEFL 240
QY      252 CKCEPKIHIDRIDKKDVEEEOEHMRWMLHERFEIKDKMLIEFYSPDERRRKRPFGKS 311
Db      241 CKCEPKIHIDRIDKKDVEEEOEHMRWMLHERFEIKDKMLIEFYSPDERRRKRPFGKS 300
QY      312 VNSKLSIKTLPBMLILSGLTAGMLMTDAGRKLVTWNTWYIGTLGCLMTWYTIKA 364
Db      301 VNSKLSIKTLPBMLILSGLTAGMLMTDAGRKLVTWNTWYIGTLGCLMTWYTIKA 353
```

```
Db      301 VNSKLSIKTLPBMLILSGLTAGMLMTDAGRKLVTWNTWYIGTLGCLMTWYTIKA 353
```

```
RESULT 10
PCT-US01-11988-1475
; Sequence 1475, Application PC/TUS0111988
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Human Genome Sciences, Inc.
```

```
; TITLE OF INVENTION: Albumin Fusion Proteins
```

```
; FILE REFERENCE: PF546PCT
```

```
; CURRENT APPLICATION NUMBER: PCT/US01/11988
```

```
; CURRENT FILING DATE: 2001-01-12
```

```
; PRIOR APPLICATION NUMBER: 60/229, 358
```

```
; PRIOR FILING DATE: 2000-04-12
```

```
; PRIOR APPLICATION NUMBER: 60/256, 931
```

```
; PRIOR FILING DATE: 2000-12-21
```

```
; PRIOR APPLICATION NUMBER: 60/199, 384
```

```
; PRIOR FILING DATE: 2000-04-25
```

```
; NUMBER OF SEQ ID NOS: 2267
```

```
; SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 1475
```

```
; LENGTH: 353
```

```
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
```

```
PCT-US01-11988-1475
```

```
Query Match          97.2%; Score 1871; DB 1; Length 353;
Best Local Similarity 100.0%; Pred. No. 1.1e-183;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      12  MRLLPSVVLGTAPTYVVLAMGWRLLSAFLPARFYQALDDRLCYCYQSQWVLEFFENYTG 71
Db      1  MRLLPSVVLGTAPTYVVLAMGWRLLSAFLPARFYQALDDRLCYCYQSQWVLEFFENYTG 60
QY      72  VQILLYGDLPKKNENIYYLANHSTVDWIADILAIROMALGHVRYLKEGLKMLPLYGC 131
Db      61  VQILLYGDLPKKNENIYYLANHSTVDWIADILAIROMALGHVRYLKEGLKMLPLYGC 120
QY      132 YFAOHGGIYKRSKAKNEKEMRNKLOSVDAGTPMYLVIFPESTRNPEQTKYLSASQAF 191
Db      121 YFAOHGGIYKRSKAKNEKEMRNKLOSVDAGTPMYLVIFPESTRNPEQTKYLSASQAF 180
QY      192 AAOAGLAVLKHLVLPRIKATHVAFDCMKNYLDIYDVTYVVEGKDDGGQRESPTMTTEFL 251
Db      181 AAOAGLAVLKHLVLPRIKATHVAFDCMKNYLDIYDVTYVVEGKDDGGQRESPTMTTEFL 240
QY      252 CKCEPKIHIDRIDKKDVEEEOEHMRWMLHERFEIKDKMLIEFYSPDERRRKRPFGKS 311
Db      241 CKCEPKIHIDRIDKKDVEEEOEHMRWMLHERFEIKDKMLIEFYSPDERRRKRPFGKS 300
QY      312 VNSKLSIKTLPBMLILSGLTAGMLMTDAGRKLVTWNTWYIGTLGCLMTWYTIKA 364
Db      301 VNSKLSIKTLPBMLILSGLTAGMLMTDAGRKLVTWNTWYIGTLGCLMTWYTIKA 353

RESULT 11
US-09-833-245-1474
; Sequence 1474, Application US/09833245
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
```

SEQ ID NO 1474  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-833-245-1474

Query Match 57.2%; Score 1871; DB 22; Length 353;  
Best Local Similarity 100.0%; Pred. No. 1,1e-183;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 MRYLLPSVLLGTAFTYVLAWGVRLSAPLPARYQALDRLKCVQSMVLFEEENTG 71  
DB 1 MRYLLPSVLLGTAFTYVLAWGVRLSAPLPARYQALDRLKCVQSMVLFEEENTG 60  
OY 72 VOILLYGDLPKKNENIYLANHSTVDWIVADILAIRONALGHVRYVLEKGLKMLPLXGC 131  
DB 61 VOILLYGDLPKKNENIYLANHSTVDWIVADILAIRONALGHVRYVLEKGLKMLPLXGC 120  
OY 132 YFAOHGGIYVRSKAFENKEKRNKLOSVDAGTPMYLVIPEEGTRYNEQTKVLSASQAF 191  
DB 121 YFAOHGGIYVRSKAFENKEKRNKLOSVDAGTPMYLVIPEEGTRYNEQTKVLSASQAF 180  
OY 192 AAOGLAVLKHVLPRIKATHAVDCMKNYLDATYDVTVYVEGKDDGQORESPPTMEFL 251  
DB 181 AAOGLAVLKHVLPRIKATHAVDCMKNYLDATYDVTVYVEGKDDGQORESPPTMEFL 240  
OY 252 CKECPKIHIDRIDKDVPEQEHMRRLHEREIKDKMLIEFYESPDPERRRRFPKGS 311  
DB 241 CKECPKIHIDRIDKDVPEQEHMRRLHEREIKDKMLIEFYESPDPERRRRFPKGS 300  
OY 312 VNSKLSIKTLPMSLLISGLTAGMLMTDAGRKLYVNTWYIGTLLGCLMWITKA 364  
DB 301 VNSKLSIKTLPMSLLISGLTAGMLMTDAGRKLYVNTWYIGTLLGCLMWITKA 353

RESULT 12  
US-09-833-245-1475  
Sequence 1475, Application US/09833245  
GENERAL INFORMATION:  
APPLICANT: Human Genome Sciences, Inc.  
TITLE OF INVENTION: Albumin Fusion Proteins  
FILE REFERENCE: P546PCT  
CURRENT APPLICATION NUMBER: US/09/833,245  
CURRENT FILING DATE: 2001-04-12  
PRIOR APPLICATION NUMBER: 60/229,358  
PRIOR FILING DATE: 2000-04-12  
PRIOR APPLICATION NUMBER: 60/256,931  
PRIOR FILING DATE: 2000-12-21  
PRIOR APPLICATION NUMBER: 60/199,384  
PRIOR FILING DATE: 2000-04-25  
NUMBER OF SEQ ID NOS: 2267  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 1475  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-833-245-1475

Query Match 97.2%; Score 1871; DB 22; Length 353;  
Best Local Similarity 100.0%; Pred. No. 1,1e-183;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 MRYLLPSVLLGTAFTYVLAWGVRLSAPLPARYQALDRLKCVQSMVLFEEENTG 71  
DB 1 MRYLLPSVLLGTAFTYVLAWGVRLSAPLPARYQALDRLKCVQSMVLFEEENTG 60  
OY 72 VOILLYGDLPKKNENIYLANHSTVDWIVADILAIRONALGHVRYVLEKGLKMLPLXGC 131  
DB 61 VOILLYGDLPKKNENIYLANHSTVDWIVADILAIRONALGHVRYVLEKGLKMLPLXGC 120  
OY 132 YFAOHGGIYVRSKAFENKEKRNKLOSVDAGTPMYLVIPEEGTRYNEQTKVLSASQAF 191

DB 121 YFAOHGGIYVRSKAFENKEKRNKLOSVDAGTPMYLVIPEEGTRYNEQTKVLSASQAF 180  
OY 192 AAOGLAVLKHVLPRIKATHAVDCMKNYLDATYDVTVYVEGKDDGQORESPPTMEFL 251  
DB 181 AAOGLAVLKHVLPRIKATHAVDCMKNYLDATYDVTVYVEGKDDGQORESPPTMEFL 240  
OY 252 CKECPKIHIDRIDKDVPEQEHMRRLHEREIKDKMLIEFYESPDPERRRRFPKGS 311  
DB 241 CKECPKIHIDRIDKDVPEQEHMRRLHEREIKDKMLIEFYESPDPERRRRFPKGS 300  
OY 312 VNSKLSIKTLPMSLLISGLTAGMLMTDAGRKLYVNTWYIGTLLGCLMWITKA 364  
DB 301 VNSKLSIKTLPMSLLISGLTAGMLMTDAGRKLYVNTWYIGTLLGCLMWITKA 353

RESULT 13  
US-09-853-526-4  
Sequence 4, Application US/09853526  
GENERAL INFORMATION:  
APPLICANT: Cohen, Daniel  
APPLICANT: Blumenfeld, Marta  
APPLICANT: Ilyia, Chumakov  
APPLICANT: Bougueleret, Lydie  
TITLE OF INVENTION: PROSTATE CANCER GENE  
FILE REFERENCE: GENSET.18CPICP  
CURRENT APPLICATION NUMBER: US/09/853,526  
CURRENT FILING DATE: 2001-05-11  
PRIOR APPLICATION NUMBER: 09/338,907  
PRIOR FILING DATE: 1999-06-23  
PRIOR APPLICATION NUMBER: 08/996,306  
PRIOR FILING DATE: 1997-12-22  
PRIOR APPLICATION NUMBER: 60/099,658  
PRIOR FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: 09/218,207  
PRIOR FILING DATE: 1998-12-22  
NUMBER OF SEQ ID NOS: 578  
SOFTWARE: Patent.lm  
SEQ ID NO 4  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 1..33  
OTHER INFORMATION: Rao and Argos identification method, potential helix  
NAME/KEY: HELIX  
LOCATION: 4..20  
OTHER INFORMATION: Klein, Kanehisa and Delisi identification method, potent  
LOCATION: 4..24  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
OTHER INFORMATION: helix  
NAME/KEY: MYRISTATE  
LOCATION: 12..16  
OTHER INFORMATION: Prosite match  
NAME/KEY: HELIX  
LOCATION: 50..70  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
OTHER INFORMATION: helix  
NAME/KEY: CARBOHYD  
LOCATION: 57..59  
OTHER INFORMATION: Prosite match  
NAME/KEY: HELIX  
LOCATION: 76..96  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 78  
OTHER INFORMATION: potential Tyrosine kinase site, Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 84  
OTHER INFORMATION: potential caseine kinase II site, Prosite match

```
NAME/KEY: SITE
LOCATION: 94..115
OTHER INFORMATION: potential Leucine zipper site, Prosite match
NAME/KEY: MYRISTATE
LOCATION: 119..123
OTHER INFORMATION: potential site, Prosite match
NAME/KEY: PHOSPHORYLATION
LOCATION: 133
OTHER INFORMATION: potential protein kinase C, Prosite match
NAME/KEY: PHOSPHORYLATION
LOCATION: 147
OTHER INFORMATION: potential caseine kinase II site, Prosite match
NAME/KEY: PHOSPHORYLATION
LOCATION: 194
OTHER INFORMATION: potential protein kinase C, Prosite match
NAME/KEY: PHOSPHORYLATION
LOCATION: 215
OTHER INFORMATION: potential Tyrosine kinase site, Prosite match
NAME/KEY: SULFATATION
LOCATION: 221
OTHER INFORMATION: Prosite match
NAME/KEY: PHOSPHORYLATION
LOCATION: 233
OTHER INFORMATION: potential cAMP and cGMP dependant protein kinase site, Prosite match
NAME/KEY: PHOSPHORYLATION
LOCATION: 235
OTHER INFORMATION: potential caseine kinase II site, Prosite match
NAME/KEY: PHOSPHORYLATION
LOCATION: 306
OTHER INFORMATION: potential protein kinase C, Prosite match
NAME/KEY: HELIX
LOCATION: 310..330
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F
NAME/KEY: MYRISTATE
LOCATION: 319..323
OTHER INFORMATION: helix
NAME/KEY: MYRISTATE
LOCATION: 323..327
OTHER INFORMATION: Prosite match
NAME/KEY: MYRISTATE
LOCATION: 329
OTHER INFORMATION: Prosite match
NAME/KEY: AMIDATION
LOCATION: 329
OTHER INFORMATION: Prosite match
NAME/KEY: HELIX
LOCATION: 333..353
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F
NAME/KEY: MYRISTATE
LOCATION: 341..345
OTHER INFORMATION: Prosite match
NAME/KEY: PHOSPHORYLATION
LOCATION: 350
OTHER INFORMATION: potential protein kinase C, Prosite match
US-09-853-526-4
```

Query Match 97.2%; Score 1871; DB 22; Length 353;  
Best Local Similarity 100.0%; Pred. No. 1,1e-183;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 12 MRLPLSPVLLGAPTYVLAAGVWRLLSAFLPARFYQALDDRLXYCYQSVMVLFEEFNNG 71
DB 1 MRLPLSPVLLGAPTYVLAAGVWRLLSAFLPARFYQALDDRLXYCYQSVMVLFEEFNNG 60
QY 72 VOLLGVDLPKKNKENTIIYLANHOSYVDWIVADILAIKRONALGHVRVLEKGLKMLPLXGC 131
DB 61 VOLLGVDLPKKNKENTIIYLANHOSYVDWIVADILAIKRONALGHVRVLEKGLKMLPLXGC 120
QY 132 YFAOHGIGIVKRSKAKNEKEMRNKLOSIVDAGTDMVLVIFPESTRNPEQTKVLSASQAF 191
DB 121 YFAOHGIGIVKRSKAKNEKEMRNKLOSIVDAGTDMVLVIFPESTRNPEQTKVLSASQAF 180
```

```
QY 192 AAGRGVLAVLKHVLPRIKATHVAFDCMKNYLAIYDVTVVYEGKDDGGORRESPTMTTEFL 251
DB 181 AAGRGVLAVLKHVLPRIKATHVAFDCMKNYLAIYDVTVVYEGKDDGGORRESPTMTTEFL 240
QY 252 CKCEPKIHIIHIDRIDKKVDPEOEHRMRWLHEHFEIKDKMLIEFYSPDERKRRPPGKS 311
DB 241 CKCEPKIHIIHIDRIDKKVDPEOEHRMRWLHEHFEIKDKMLIEFYSPDERKRRPPGKS 300
QY 312 VNSKLSIKKTPSLMILISGLTAGMLMTDAGRKLYVTMTWITGTLGCLMTYTIKA 364
DB 301 VNSKLSIKKTPSLMILISGLTAGMLMTDAGRKLYVTMTWITGTLGCLMTYTIKA 353
```

RESULT 14  
US-09-901-484A-4

```
Sequence 4, Application US/09901484A
GENERAL INFORMATION:
APPLICANT: Cohen, Daniel
APPLICANT: Blumenfeld, Marta
APPLICANT: Chumakov, Ilya
APPLICANT: Bougueleret, Lydie
TITLE OF INVENTION: Prostate Cancer Gene
FILE REFERENCE: GEN-T11XC3D2
CURRENT APPLICATION NUMBER: US/09/901,484A
PRIOR FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: US 08/996,306
PRIOR FILING DATE: 1997-12-22
PRIOR APPLICATION NUMBER: US 60/099,658
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: US 09/218,207
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: US 09/338,907
PRIOR FILING DATE: 1999-06-23
PRIOR APPLICATION NUMBER: US 09/853,526
PRIOR FILING DATE: 2001-05-11
NUMBER OF SEQ ID NOS: 578
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 353
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: HELIX
LOCATION: (1)..(33)
OTHER INFORMATION: Rao and Argos identification method, potential helix
NAME/KEY: HELIX
LOCATION: (4)..(20)
OTHER INFORMATION: Klein, Kanehisa and Delisi identification method, potential he
NAME/KEY: HELIX
LOCATION: (4)..(24)
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
OTHER INFORMATION: potential helix
NAME/KEY: LIPID
LOCATION: (12)..(16)
OTHER INFORMATION: MYRISTATE, Prosite match
NAME/KEY: HELIX
LOCATION: (50)..(70)
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
OTHER INFORMATION: potential helix
NAME/KEY: CARBOHYD
LOCATION: (57)..(59)
OTHER INFORMATION: Prosite match
NAME/KEY: HELIX
LOCATION: (76)..(96)
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
OTHER INFORMATION: potential helix
NAME/KEY: MOD RES
LOCATION: (78)..(78)
OTHER INFORMATION: PHOSPHORYLATION, potential Tyrosine kinase site, Prosite match
NAME/KEY: MOD RES
LOCATION: (84)..(84)
OTHER INFORMATION: PHOSPHORYLATION, potential caseine kinase II site, Prosite mat
NAME/KEY: SITE
```

```
LOCATION: (94)..(115)
OTHER INFORMATION: Potential Leucine Zipper site, Prosite match
NAME/KEY: LIPID
LOCATION: (119)..(123)
OTHER INFORMATION: MYR::STATE, Prosite match
NAME/KEY: MOD_RES
LOCATION: (133)..(133)
OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match
NAME/KEY: MOD_RES
LOCATION: (147)..(147)
OTHER INFORMATION: PHOSPHORYLATION, potential casein kinase II, Prosite match
NAME/KEY: MOD_RES
LOCATION: (194)..(194)
OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match
NAME/KEY: MOD_RES
LOCATION: (215)..(215)
OTHER INFORMATION: PHOSPHORYLATION, potential Tyrosine kinase site, Prosite match
NAME/KEY: MOD_RES
LOCATION: (221)..(221)
OTHER INFORMATION: SULFATATION, Prosite match
NAME/KEY: MOD_RES
LOCATION: (233)..(233)
OTHER INFORMATION: PHOSPHORYLATION, potential cAMP and cGMP dependant protein kinase
NAME/KEY: MOD_RES
LOCATION: (235)..(235)
OTHER INFORMATION: PHOSPHORYLATION, potential casein kinase II site, Prosite match
NAME/KEY: MOD_RES
LOCATION: (306)..(306)
OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match
NAME/KEY: HELIX
LOCATION: (310)..(330)
OTHER INFORMATION: Eisenberg, Schwarz, Komaromy, Wall identification method,
NAME/KEY: LIPID
LOCATION: (319)..(323)
OTHER INFORMATION: MYRISTATE, Prosite match
NAME/KEY: LIPID
LOCATION: (323)..(327)
OTHER INFORMATION: MYRISTATE, Prosite match
NAME/KEY: MOD_RES
LOCATION: (329)..(329)
OTHER INFORMATION: AMIDATION, Prosite match
NAME/KEY: HELIX
LOCATION: (333)..(353)
OTHER INFORMATION: Eisenberg, Schwarz, Komaromy, Wall identification method,
NAME/KEY: LIPID
LOCATION: (341)..(345)
OTHER INFORMATION: MYRISTATE, Prosite match
NAME/KEY: MOD_RES
LOCATION: (350)..(350)
OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match
US-09-901-484A-4
```

Query Match 97.2%: Score 1871; DB 23; Length 353;  
Best Local Similarity 100.0%; Pred. No. 1.1e-183;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 12 MRYLLPSVLLGTAPTYVLAMGVRRLLSAFLPARFYQALDRLCYQSVLFFFEYNTG 71
DB 1 MRYLLPSVLLGTAPTYVLAMGVRRLLSAFLPARFYQALDRLCYQSVLFFFEYNTG 60
QY 72 VOILLGDLPRKNENIYYLANHOSTVDMIVADILAIQNALGHRVYLKEGLKMLPLYGC 131
DB 61 VOILLGDLPRKNENIYYLANHOSTVDMIVADILAIQNALGHRVYLKEGLKMLPLYGC 120
QY 132 YFAOHGIIYVRSKAFNEKEIRNKLQSYVDAGTPMYLVIFPEGTRYNPEQTKVLSASQAF 191
DB 121 YFAOHGIIYVRSKAFNEKEIRNKLQSYVDAGTPMYLVIFPEGTRYNPEQTKVLSASQAF 180
QY 192 AAQGLAVLKHVLTPIRKATIHVAFDCMKNYLDIAYDVTVVYEGKDDGQRRSEPTMTBEFL 251
```

```
DB 181 AAQGLAVLKHVLTPIRKATIHVAFDCMKNYLDIAYDVTVVYEGKDDGQRRSEPTMTBEFL 240
QY 252 CKECPKIHIDRIDDKDVPEOEHRMRWMLHEREIKDKMLIEFYESPDPERRRRPFGKS 311
DB 241 CKECPKIHIDRIDDKDVPEOEHRMRWMLHEREIKDKMLIEFYESPDPERRRRPFGKS 300
QY 312 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTNWIYGTLLGCLMWTIKA 364
DB 301 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTNWIYGTLLGCLMWTIKA 353

RESULT 15
US-09-915-582-59
Sequence 59, Application US/09915582
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OR INVENTION: 17 Human Secreted Proteins
FILE REFERENCE: P5723P1
CURRENT APPLICATION NUMBER: US/09/915,582
PRIOR FILING DATE: 2001-07-27
PRIOR APPLICATION NUMBER: PCT/US01/01431
PRIOR FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: 60/179,065
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/180,628
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: 60/231,968
PRIOR FILING DATE: 2000-09-12
NUMBER OF SEQ ID NOS: 97
SOFTWARE: Patentl Ver. 2.0
SEQ ID NO 59
LENGTH: 353
TYPE: PRT
ORGANISM: Homo sapiens
US-09-915-582-59
```

Query Match 97.2%: Score 1871; DB 23; Length 353;  
Best Local Similarity 100.0%; Pred. No. 1.1e-183;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 12 MRYLLPSVLLGTAPTYVLAMGVRRLLSAFLPARFYQALDRLCYQSVLFFFEYNTG 71
DB 1 MRYLLPSVLLGTAPTYVLAMGVRRLLSAFLPARFYQALDRLCYQSVLFFFEYNTG 60
QY 72 VOILLGDLPRKNENIYYLANHOSTVDMIVADILAIQNALGHRVYLKEGLKMLPLYGC 131
DB 61 VOILLGDLPRKNENIYYLANHOSTVDMIVADILAIQNALGHRVYLKEGLKMLPLYGC 120
QY 132 YFAOHGIIYVRSKAFNEKEIRNKLQSYVDAGTPMYLVIFPEGTRYNPEQTKVLSASQAF 191
DB 121 YFAOHGIIYVRSKAFNEKEIRNKLQSYVDAGTPMYLVIFPEGTRYNPEQTKVLSASQAF 180
QY 192 AAQGLAVLKHVLTPIRKATIHVAFDCMKNYLDIAYDVTVVYEGKDDGQRRSEPTMTBEFL 251
DB 181 AAQGLAVLKHVLTPIRKATIHVAFDCMKNYLDIAYDVTVVYEGKDDGQRRSEPTMTBEFL 240
QY 252 CKECPKIHIDRIDDKDVPEOEHRMRWMLHEREIKDKMLIEFYESPDPERRRRPFGKS 311
DB 241 CKECPKIHIDRIDDKDVPEOEHRMRWMLHEREIKDKMLIEFYESPDPERRRRPFGKS 300
QY 312 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTNWIYGTLLGCLMWTIKA 364
DB 301 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTNWIYGTLLGCLMWTIKA 353
```

Search completed: August 28, 2002, 11:14:11  
Job time: 490 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:15:41 ; Search time 84.08 Seconds

(without alignments)  
1047.889 Million cell updates/sec

Title: US-09-853-526-5

Perfect score: 1924  
Sequence: 1 MLTSLVLTHTYSMRYLPSVY.....YVNTWIVGTLLGLMTWTKA 364

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747981 seqs, 242050750 residues

Total number of hits satisfying chosen parameters: 747981

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Pending\_Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1924	100.0	364	US-10-184-648-63	Sequence 63, Appl
2	1868	97.1	353	US-09-629-469A-13028	Sequence 13028, A
3	1427	74.2	269	US-10-074-045-47	Sequence 47, Appl
4	390.5	20.3	375	US-09-935-625-11061	Sequence 11061, A
5	390.5	20.3	375	US-09-935-625-14197	Sequence 14197, A
6	390.5	20.3	375	US-09-935-625-14201	Sequence 14201, A
7	390.5	20.3	375	US-09-935-625-14254	Sequence 14254, A
8	390.5	20.3	375	US-09-935-625-23009	Sequence 23009, A
9	390.5	20.3	375	US-09-935-625-23025	Sequence 23025, A
10	387.5	20.1	351	US-09-935-625-11062	Sequence 11062, A
11	387.5	20.1	351	US-09-935-625-14198	Sequence 14198, A
12	387.5	20.1	351	US-09-935-625-14202	Sequence 14202, A
13	387.5	20.1	351	US-09-935-625-14255	Sequence 14255, A
14	387.5	20.1	351	US-09-935-625-23010	Sequence 23010, A
15	387.5	20.1	351	US-09-935-625-23026	Sequence 23026, A
16	387.5	20.1	375	US-09-935-625-11471	Sequence 11471, A
17	384.5	20.0	351	US-09-935-625-11472	Sequence 11472, A
18	381	19.8	343	US-09-935-625-11063	Sequence 11063, A
19	381	19.8	343	US-09-935-625-11473	Sequence 11473, A
20	381	19.8	343	US-09-935-625-11474	Sequence 11474, A
21	381	19.8	343	US-09-935-625-14199	Sequence 14199, A
22	381	19.8	343	US-09-935-625-14203	Sequence 14203, A
23	381	19.8	343	US-09-935-625-14256	Sequence 14256, A
24	381	19.8	343	US-09-935-625-23011	Sequence 23011, A
25	374.5	19.5	400	US-09-935-625-23027	Sequence 23027, A
26	305.5	15.9	414	US-10-121-062-314	Sequence 314, App

27	305.5	15.9	414	6	US-10-063-502-102	Sequence 102, App
28	305.5	15.9	414	6	US-10-063-510-102	Sequence 102, App
29	305.5	15.9	414	6	US-10-063-512-102	Sequence 102, App
30	305.5	15.9	414	6	US-10-063-513-102	Sequence 102, App
31	305.5	15.9	414	6	US-10-063-514-102	Sequence 102, App
32	305.5	15.9	414	6	US-10-063-515-102	Sequence 102, App
33	305.5	15.9	414	6	US-10-063-516-102	Sequence 102, App
34	305.5	15.9	414	6	US-10-063-517-102	Sequence 102, App
35	305.5	15.9	414	6	US-10-063-518-102	Sequence 102, App
36	305.5	15.9	414	6	US-10-063-519-102	Sequence 102, App
37	305.5	15.9	414	6	US-10-063-520-102	Sequence 102, App
38	305.5	15.9	414	6	US-10-063-521-102	Sequence 102, App
39	305.5	15.9	414	6	US-10-063-522-102	Sequence 102, App
40	305.5	15.9	414	6	US-10-063-523-102	Sequence 102, App
41	305.5	15.9	414	6	US-10-063-524-102	Sequence 102, App
42	305.5	15.9	414	6	US-10-063-526-102	Sequence 102, App
43	305.5	15.9	414	6	US-10-063-527-102	Sequence 102, App
44	305.5	15.9	414	6	US-10-063-528-102	Sequence 102, App
45	305.5	15.9	414	6	US-10-063-529-102	Sequence 102, App

#### ALIGNMENTS

RESULT 1  
US-10-184-648-63  
Sequence 63, Application US/10184648  
GENERAL INFORMATION:  
APPLICANT: Meyers, Rachel E.  
APPLICANT: Williamson, Mark  
APPLICANT: Tsai, Fong-Ying  
APPLICANT: Hunter, John J.  
APPLICANT: Macbeth, Kyle J.  
APPLICANT: Rudolph-Owen, Laura A.  
APPLICANT: Leiby, Kevin R.  
APPLICANT: Kapeller-Libermann, Rosana  
APPLICANT: Olandt, Peter J.  
TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE FAMILY MEMBERS AND USES THEREOF  
FILE REFERENCE: 10448-192001  
CURRENT APPLICATION NUMBER: US/10/184, 648  
CURRENT FILING DATE: 2002-06-27  
PRIOR APPLICATION NUMBER: US 09/815, 028  
PRIOR FILING DATE: 2001-03-22  
PRIOR APPLICATION NUMBER: PCT/US01/09358  
PRIOR FILING DATE: 2001-03-22  
PRIOR APPLICATION NUMBER: US 60/191, 964  
PRIOR FILING DATE: 2000-03-24  
PRIOR APPLICATION NUMBER: US 09/801, 220  
PRIOR FILING DATE: 2001-03-07  
PRIOR APPLICATION NUMBER: PCT/US01/07269  
PRIOR FILING DATE: 2001-03-07  
PRIOR APPLICATION NUMBER: US 60/187, 456  
PRIOR FILING DATE: 2000-03-07  
PRIOR APPLICATION NUMBER: US 09/816, 714  
PRIOR FILING DATE: 2001-03-23  
PRIOR APPLICATION NUMBER: PCT/US01/09468  
PRIOR FILING DATE: 2001-03-23  
PRIOR APPLICATION NUMBER: US 60/191, 865  
PRIOR FILING DATE: 2000-03-24  
PRIOR APPLICATION NUMBER: US 09/844, 948  
PRIOR FILING DATE: 2001-04-27  
PRIOR APPLICATION NUMBER: PCT/US01/13805  
PRIOR FILING DATE: 2001-04-27  
PRIOR APPLICATION NUMBER: US 60/200, 604  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: US 09/861, 164  
PRIOR FILING DATE: 2001-05-18  
PRIOR APPLICATION NUMBER: PCT/US01/16292  
PRIOR FILING DATE: 2001-05-18  
PRIOR APPLICATION NUMBER: US 60/205, 408  
PRIOR FILING DATE: 2000-05-19  
PRIOR APPLICATION NUMBER: US 09/883, 060  
PRIOR FILING DATE: 2001-06-15

```

: PRIOR APPLICATION NUMBER: PCT/US01/19138
: PRIOR FILING DATE: 2001-06-15
: PRIOR APPLICATION NUMBER: US 60/212,079
: PRIOR FILING DATE: 2000-06-15
: PRIOR APPLICATION NUMBER: US 09/962,678
: PRIOR FILING DATE: 2001-09-25
: PRIOR APPLICATION NUMBER: PCT/US01/29963
: PRIOR FILING DATE: 2001-09-25
: PRIOR APPLICATION NUMBER: US 60/235,044
: PRIOR FILING DATE: 2000-09-25
: PRIOR APPLICATION NUMBER: US 09/973,457
: PRIOR FILING DATE: 2001-10-09
: PRIOR APPLICATION NUMBER: US 60/238,849
: PRIOR FILING DATE: 2000-10-06
: PRIOR APPLICATION NUMBER: US 10/072,285
: PRIOR FILING DATE: 2002-02-08
: PRIOR APPLICATION NUMBER: PCT/US02/03736
: PRIOR FILING DATE: 2002-02-08
: PRIOR APPLICATION NUMBER: US 60/267,494
: PRIOR FILING DATE: 2001-02-08
: PRIOR APPLICATION NUMBER: US 09/817,910
: PRIOR FILING DATE: 2001-03-26
: PRIOR APPLICATION NUMBER: PCT/US01/09633
: PRIOR FILING DATE: 2001-03-26
: PRIOR APPLICATION NUMBER: US 60/192,092
: PRIOR FILING DATE: 2000-03-24
: PRIOR APPLICATION NUMBER: US 09/842,528
: PRIOR FILING DATE: 2001-04-25
: PRIOR APPLICATION NUMBER: PCT/US01/40607
: PRIOR FILING DATE: 2001-04-25
: PRIOR APPLICATION NUMBER: US 60/199,500
: PRIOR FILING DATE: 2000-04-25
: PRIOR APPLICATION NUMBER: US 09/882,836
: PRIOR FILING DATE: 2001-06-15
: PRIOR APPLICATION NUMBER: PCT/US01/19543
: PRIOR FILING DATE: 2001-06-15
: PRIOR APPLICATION NUMBER: US 60/211,730
: PRIOR FILING DATE: 2000-06-15
: PRIOR APPLICATION NUMBER: US 09/882,872
: PRIOR FILING DATE: 2001-06-15
: PRIOR APPLICATION NUMBER: PCT/US01/19153
: PRIOR FILING DATE: 2001-06-15
: PRIOR APPLICATION NUMBER: US 60/212,077
: PRIOR FILING DATE: 2000-06-15
: NUMBER OF SEQ ID NOS: 90
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 63
: LENGTH: 364
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-184-648-63

```

```

Query Match          100.0%; Score 1924; DB 6; Length 364;
Best Local Similarity 100.0%; Pred. No. 1.5e-171;
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLTSLVHTYSMRYLPSVVLGTAPTYVLAMGWRLSAPLPARFYQALDDRLCYVQOS 60
DB 1 MLTSLVHTYSMRYLPSVVLGTAPTYVLAMGWRLSAPLPARFYQALDDRLCYVQOS 60
QY 61 MLEFFENYTGVOILLYGDLP:KNENITYLANHSTVDMIVADILAIKRONALGHVRYVLK 120
DB 61 MLEFFENYTGVOILLYGDLP:KNENITYLANHSTVDMIVADILAIKRONALGHVRYVLK 120
QY 121 EGLKWLPLYGCFYQHGIGIYKRSKAKNEKEMRKLDQSYVDAGTPMYLVIFPEGTRYNPE 180
DB 121 EGLKWLPLYGCFYQHGIGIYKRSKAKNEKEMRKLDQSYVDAGTPMYLVIFPEGTRYNPE 180
QY 181 OTKVLASQAFPAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDALIVDTVYVEGKDDGGO 240
DB 181 OTKVLASQAFPAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDALIVDTVYVEGKDDGGO 240

```

```

QY 241 RRESPTMTFELCKECPKIHIDRIDKKVPEEOEHMRWHERFEIKDKMLIEFYSPD 300
DB 241 RRESPTMTFELCKECPKIHIDRIDKKVPEEOEHMRWHERFEIKDKMLIEFYSPD 300
QY 301 PERKRRPFGKSVNSKLSIKTKLPMSLITSLTACMLMMDACRKLYVMTWIYGTLLGCLMV 360
DB 301 PERKRRPFGKSVNSKLSIKTKLPMSLITSLTACMLMMDACRKLYVMTWIYGTLLGCLMV 360
QY 361 TIKKA 364
DB 361 TIKKA 364

```

```

RESULT 2
US-09-629-469A-13028
: Sequence 13028, Application us/09629469A
: GENERAL INFORMATION:
: APPLICANT: OTA, TOSHIO
: APPLICANT: ISOGAI, TAKAO
: APPLICANT: NISHIKAWA, TETSUO
: APPLICANT: HAYASHI, KOJI
: APPLICANT: SATTO, KAORI
: APPLICANT: YAMAMOTO, JUNICHI
: APPLICANT: ISHII, SHIZUKO
: APPLICANT: SUGIYAMA, TOMOYASU
: APPLICANT: WAKAMATSU, AI
: APPLICANT: NAGAI, KEIICHI
: APPLICANT: OTSUKI, TETSUJI
: TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE
: FILE REFERENCE: 084335/0123
: CURRENT APPLICATION NUMBER: US/09/629,469A
: PRIOR FILING DATE: 2000-07-28
: PRIOR APPLICATION NUMBER: JP 1999-248036
: PRIOR FILING DATE: 1999-07-29
: PRIOR APPLICATION NUMBER: JP 1999-300253
: PRIOR FILING DATE: 1999-08-27
: PRIOR APPLICATION NUMBER: JP 2000-118776
: PRIOR FILING DATE: 2000-01-11
: PRIOR APPLICATION NUMBER: JP 2000-183767
: PRIOR FILING DATE: 2000-05-02
: PRIOR APPLICATION NUMBER: JP 2000-241899
: PRIOR FILING DATE: 2000-06-09
: PRIOR APPLICATION NUMBER: 60/159,590
: PRIOR FILING DATE: 1999-10-18
: PRIOR APPLICATION NUMBER: 60/183,322
: PRIOR FILING DATE: 2000-02-17
: NUMBER OF SEQ ID NOS: 19025
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 13028
: LENGTH: 353
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-629-469A-13028

```

```

Query Match          97.1%; Score 1866; DB 5; Length 353;
Best Local Similarity 99.7%; Pred. No. 2.5e-166;
Matches 352; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 12 MRYLLPSVVLGTAPTYVLAMGWRLSAPLPARFYQALDDRLCYVQOSWVLPFFENYTG 71
DB 1 MRYLLPSVVLGTAPTYVLAMGWRLSAPLPARFYQALDDRLCYVQOSWVLPFFENYTG 71
QY 72 VOILLYGDLP:KNENITYLANHSTVDMIVADILAIKRONALGHVRYVLKGLKMLPLYGC 131
DB 72 VOILLYGDLP:KNENITYLANHSTVDMIVADILAIKRONALGHVRYVLKGLKMLPLYGC 131
QY 132 YFAOHGIGIYKRSKAKNEKEMRKLDQSYVDAGTPMYLVIFPEGTRYNPEQTKVLASQAF 191
DB 132 YFAOHGIGIYKRSKAKNEKEMRKLDQSYVDAGTPMYLVIFPEGTRYNPEQTKVLASQAF 191
QY 191 YFAOHGIGIYKRSKAKNEKEMRKLDQSYVDAGTPMYLVIFPEGTRYNPEQTKVLASQAF 180
DB 191 YFAOHGIGIYKRSKAKNEKEMRKLDQSYVDAGTPMYLVIFPEGTRYNPEQTKVLASQAF 180
QY 251 AAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDALIVDTVYVEGKDDGGOORRESPTMTFEL 251
DB 251 AAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDALIVDTVYVEGKDDGGOORRESPTMTFEL 251

```



```
Db 181 AAGRGIAVLKHLVLPRIKATHVAFDCMKNYLDAIDVTYVVEEGKDDGGQRRSPTMTFEL 240
QY 252 CKECPKIHIDIRIDKDDPEOEHRMRMLHEFEIKDKMLIEFYSPDERRRKRRPGKS 311
Db 241 CKECPKIHIDIRIDKDDPEOEHRMRMLHEFEIKDKMLIEFYSPDERRRKRRPGKS 300
QY 312 VNSKLSIKKTLPMSMLILSGITAGMLMTDAGRKLYVTWYITGLGCLMTWYIKA 364
Db 301 VNSKLSIKKTLPMSMLILSGITAGMLMTDAGRKLYVTWYITGLGCLMTWYIKA 353

RESULT 3
US-10-074-045-47
; Sequence 47, Application US/10074045
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT2121
; CURRENT APPLICATION NUMBER: US/10/074,045
; CURRENT FILING DATE: 2002-02-14
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 47
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-074-045-47
```

```
Query Match 74.2%; Score 1427; DB 6; Length 269;
Best Local Similarity 100.0%; Pred. No. 3.3e-125;
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 96 TYDMYADILAIROMALGHVRYVLEKGLKWLPLGYCYFAOHGSIYVRSKAFNEKEMRNK 155
Db 1 TYDMYADILAIROMALGHVRYVLEKGLKWLPLGYCYFAOHGSIYVRSKAFNEKEMRNK 60
QY 156 LOSYVAGTPTMYLTPPEGTRVNPBOTKYLSSQAFRAOGLAVLKHVLPRIKATHVAF 215
Db 61 LOSYVAGTPTMYLTPPEGTRVNPBOTKYLSSQAFRAOGLAVLKHVLPRIKATHVAF 120
QY 216 DCMKNYLDAIDVTYVVEEGKDDGGQRRSPTMTFELCKECPKIHIDIRIDKDDPEOE 275
Db 121 DCMKNYLDAIDVTYVVEEGKDDGGQRRSPTMTFELCKECPKIHIDIRIDKDDPEOE 180
QY 276 HMRMLHEFEIKDKMLIEFYSPDERRRKRRPGKS VNSKLSIKKTLPMSMLILSGITAGM 335
Db 181 HMRMLHEFEIKDKMLIEFYSPDERRRKRRPGKS VNSKLSIKKTLPMSMLILSGITAGM 240
QY 336 LMTDAGRKLYVTWYITGLGCLMTWYIKA 364
Db 241 LMTDAGRKLYVTWYITGLGCLMTWYIKA 269
```

```
RESULT 4
US-09-935-625-11061
; Sequence 11061, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 11061
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..375
```

```
; OTHER INFORMATION: Ceres Seq. ID no. 1394370
US-09-935-625-11061
```

```
Query Match 20.3%; Score 390.5; DB 5; Length 375;
Best Local Similarity 26.9%; Pred. No. 5.1e-28;
Matches 101; Conservative 78; Mismatches 128; Indels 69; Gaps 9;
```

```
QY 12 MRYLPSVLLGTAPTYVLAMG-----VWRLSAPLPARYQALDRLCYQSMVLEFF 66
Db 18 LRGIICLMTVLVSTAFMMLIFWGLSAVYLRFS---IRYSRCVSEFFSGMLALMPFLF 73
QY 67 ENTGVOILLYGDLPRKKNKIYLANHSTQVDMIVADILAIROMALGHVRYVLEKGLKWL 126
Db 74 EKINKTKVITSSGDKVPCEDEVLLIANHRTVDMMYFMDLAKRGQIGNITVYVLSLMLK 133
QY 127 PLYGCFYFAOHGSIYVRSKAFNEKEMRNKLOSYVDACTPTMYLTPPEGTRVNPBOTKYL 186
Db 134 PLFGMAFHLFEFTIPERBWEDEANLQIVSFPDPRDALMLALFEGTDYTAQCO--- 190
QY 187 ASQAFRAOGLAVLKHVLPRIKATHVAFDCMKNYLDAIDVTYVVEEGKDDGGQRRSPT 246
Db 191 RSKKFAENGLPILNNVLPRTGPFVSCLOELCSLDAYVDTYIGKTR----- 239
QY 247 MTEFLCKECP-----KIHIDIRIDKDDPEOEHRMRMLHEFEIKDKMLIE 294
Db 240 -----CPSELDNNGYIPSEVHHIRININTQIPNOKEDINAMLMNTFOLKDQLND 291
QY 295 FYESPDERRRKRRPGKS VNSKLSIKKTLPMSMLILSGIT-----AGMLMTDAGRKLY 345
Db 292 FYSN-----GHPNPECTEKEFMTKKYLINCLAVIATTTICHTLFFSSIMWF---RIV 341
QY 346 VNTWYITGLGCLMTWY 361
Db 342 VS-----LACYLT 350
```

```
RESULT 5
US-09-935-625-14197
; Sequence 14197, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 14197
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..375
; OTHER INFORMATION: Ceres Seq. ID no. 3441314
US-09-935-625-14197
```

```
Query Match 20.3%; Score 390.5; DB 5; Length 375;
Best Local Similarity 26.9%; Pred. No. 5.1e-28;
Matches 101; Conservative 78; Mismatches 128; Indels 69; Gaps 9;
```

```
QY 12 MRYLPSVLLGTAPTYVLAMG-----VWRLSAPLPARYQALDRLCYQSMVLEFF 66
Db 18 LRGIICLMTVLVSTAFMMLIFWGLSAVYLRFS---IRYSRCVSEFFSGMLALMPFLF 73
QY 67 ENTGVOILLYGDLPRKKNKIYLANHSTQVDMIVADILAIROMALGHVRYVLEKGLKWL 126
Db 74 EKINKTKVITSSGDKVPCEDEVLLIANHRTVDMMYFMDLAKRGQIGNITVYVLSLMLK 133
QY 127 PLYGCFYFAOHGSIYVRSKAFNEKEMRNKLOSYVDACTPTMYLTPPEGTRVNPBOTKYL 186
```

```

Db      134 PLFGMAFLHFEFI EVERMEVDENLQIVSSFKDPRDALMLALFPEGTDYTEAKQ--- 190
QY      187 ASQAFANQAGLAVIKHVLTPRIKATNHAFCDMKMYLDALIDYVYVYEGKDD3GQGRSPT 246
        191 RSKFPAENGLPIILNVLPTPTKGFSCLQELSCSLDVAVDYVITGYTR----- 239
Db      247 MTEFLKCECP-----KIHIRIDRDKDVPDEQEHMRWMLHERFEIKOMLIE 294
        240 -----CPFLDNVYGIERSEVHHIRIRNLQIINOEDIANMLNFTQKLDQLLND 291
QY      295 FYESDPERRKRFRPKSVNSKLSIKTKLPSPMLISGLT-----AGMLMDAGKRLY 345
        292 FYSN-----GHFNEGTEKEFEFNKKYLINCLAVIAFTTICTHLPFSMWI---RIV 341
Db      346 VNTMYIGTLGCLMWT 361
Db      342 VS-----LACVYLT 350

```

RESULT 6  
 US-09-935-625-14201  
 : Sequence 14201, Application US/09935625  
 : GENERAL INFORMATION:  
 : APPLICANT: N. ALEXANDROV et al.  
 : TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE  
 : OF INVENTION: MODULATING VARIOUS RESPONSES  
 : FILE REFERENCE: 2750-1481P  
 : CURRENT APPLICATION NUMBER: US/09/935,625  
 : CURRENT FILING DATE: 2001-08-24  
 : NUMBER OF SEQ ID NOS: 33136  
 : SEQ ID NO 14201  
 : LENGTH: 375  
 : TYPE: PRT  
 : ORGANISM: Arabidopsis thaliana  
 : FEATURE:  
 : NAME/KEY: peptide  
 : LOCATION: 1..375  
 : OTHER INFORMATION: Ceres Seq. ID no. 3441318  
 : US-09-935-625-14201

	Query Match	20.3%	Score 390.5;	DB 5;	Length 375;
	Best Local Similarity	26.9%	Pred. No. 5,1e-28;		
	Matches 101; Conservative	78;	Mismatches 128;	Indels 69;	Gaps 9;
OY	12 MRLLPVSVLGTAPTYLAWG----	VWRLSLAEIPARFYQALDRLCYVQSNVLEFF	66		
Db	18 LRGICILAVLVSTAFMMLIFMGELSAVVLRFS-----	IIRSKVCYFFFGSMLAMPFLF	73		
OY	67 ENTYGVOLLVGDLPRKKNENIYYLANHOSTVMIVADIIATIQMLGNHRYLYKEGLXKL	126			
Db	74 EKINKTKTYSDDKYPCEDRYLLLNHRKEVDMMTFMDIALKKGOIGNKIYLKSSLSIML	133			
OY	127 PLVGYCFYAHOHGIIYVKRSAKFNKEEMRNKLOSUYVAGTPMYLVIFPEGRYRNPEDQTUYS	186			
Db	134 PLFGWAFLHEFIPIVERMEVDEANLRQIVSEFKDRPDALMALREPGTDYTEAKCQ---	190			
OY	187 ASQAAPAAQRGLAVLKHYLTFRPKATHVAFDCKNKTLDLIYDTVYVEGDGGORRESPT	246			
Db	191 RSKRKPAALNGPLPINNVLLPTKGFVSCLOELTSISIDAVYDVITGYKTR-----	239			
OY	247 MTEFLCKECP-----	KIHICHIDRIADKKVPVEOEIMRRWLHERFEIKDKMLE	294		
Db	240 -----	CPSFLDNVYGIEPSEVNIHNRILTLQIPNEKERINAMLMNTPOLKDQLND	291		
OY	295 FYESFDPERRRRKFPKSVANSKIUSIKTKLPSPMLISGLT-----	ACMLMTDAGRKYL	345		
Db	292 FYSN-----	GHPFNEGTEKEFTNKKYILNCIAVIATAFTICTHLTFESSMIWF----	RIV	341	
OY	346 VNTWTYIGTLGLAWT	361			
Db	342 VS-----	LACVYLTT	350		

```

RESULT 7
US-09-935-625-14254
: Sequence 14254, Application US/09935625
: GENERAL INFORMATION:
: APPLICANT: N. ALEXANDROV et al.
: TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE OF INDUCING A SPECIFIC IMMUNE RESPONSE
: TITLE OF INVENTION: MODULATING VARIOUS RESPONSES
: FILE REFERENCE: 2750-1481P
: CURRENT APPLICATION NUMBER: US/09/935,625
: CURRENT FILING DATE: 2001-08-24
: NUMBER OF SEQ. ID NOS: 33136
: SEQ ID NO 14254
: LENGTH: 375
: TYPE: PRT
: ORGANISM: Arabidopsis thaliana
: FEATURE:
: NAME/KEY: peptide
: LOCATION: 1..375
: OTHER INFORMATION: Ceres Seq. ID no. 3443248
US-09-935-625-14254

```

Query Match	20.3%	Score 390.5	DB 5	Length 375
Best Local Similarity	26.9%	Pred. No. 5.1e-28		
Matches 101; Conservative	78;	Mismatches 128;	Indels 69;	Gaps 9

```

0Y 12 MRYLLSPVLLGTAFTAYYLAWG-----VMRLLSFLPARFYQALDDRLCYQOSMVLFEFF 66
0Y 18 LRGIIICLAVLVSTANMMLIFMGFLSAVVLRLFS-----IRSRKCVSPFFSGMLMLPPELF 73
0Y 67 ENYTGVOILLYGDDEPKRKNENTIYLANHOSTVDWADILAIROMLGHVRYVLEKGLKWL 126
0Y 74 EKINTKVIYFGDGKPCEDRDYLLIANHRTVEDMMVFQMLALRKQIGIKRYVLKSLMKL 133
0Y 127 PLYGCFYFHOGHGIIYKRBAKENEKEMRKKLOSVDAGIPMLVIFPEBSTRNRPQTKYLS 186
0Y 134 PLEGFANILFEFIPERHEWDEANLRQIVSSFKDPRDALMLALFEPTDYTEAKCO--- 190
0Y 187 ASQAPAAORGLAVLEKHNVLPRIKATHVAFDCKMKNLYADIYVTVYVEEGKDDGGORRESPT 246
0Y 191 RSKRAAENGLPILNNVLPRPTKGVSCLOELSCSLDVAIVDTYIGYKTR----- 239
0Y 247 MTEFLCKECP-----KIHIIHIDRIDKDVPEOEGBMRMRMLHEREFEIKDKMLIE 294
0Y 240 -----CPSFLNDVNGIEPSEVHHIIRIRNLNQIPNOEKDINAMLNPTQLDKOLLND 291
0Y 295 FYESPDPERRRRKPPCKSVNSKLSIKTKLPSMLILSGLT-----AGMLMTDAGRRLY 345
0Y 292 FYSN-----GHFPEGTEKEKEFNTRKYLINCLAVIAFTTICTHILTFESSMWTF---RIY 341
0Y 346 VNTWIVYGLGLCLAWT 361
0Y 342 VS-----LACVYLIT 350

```

```

RESULT      8
US-09-935-625-23009
: Sequence: 23009, Application US/09935625
:
: GENERAL INFORMATION:
: APPLICANT: N. ALEXANDROV et al.
: TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE OF INDUCING A SPECIFIC IMMUNE RESPONSE
: TITLE OF INVENTION: MODULATING VARIOUS RESPONSES
: FILE REFERENCE: 2750-1481P
: CURRENT APPLICATION NUMBER: US/09/935,625
: CURRENT FILING DATE: 2001-08-24
: NUMBER OF SEQ ID NOS: 31136
: SEQ ID NO 23009
:
: LENGTH: 375
: TYPE: PRT
: ORGANISM: Arabidopsis thaliana
: FEATURE:
: NAME/KEY: peptide

```

```
; LOCATION: 1..375
; OTHER INFORMATION: Ceres Seq. ID no. 3441318
US-09-935-625-23009

Query Match          20.3%; Score 390.5; DB 5; Length 375;
Best Local Similarity 26.9%; Pred. No. 5.1e-28;
Matches 101; Conservative 78; Mismatches 128; Indels 69; Gaps 9;

QY 12 MRLTSPVLLGTAFTYVLAMG-----VWRLSAFLPAREYQALDRLCYQOSMVLFEFF 66
DB 18 LRGIILAMVLTAFMMLIFWGFLSAVVLRLFS---IRSRKCVSFFGSMWALMPFLF 73
QY 67 ENTGVOILLGDLPRKKNIIYLANHOSVVDWIVADILAIRONALGHVRYVLEKGLKWL 126
DB 74 EKINTKRVITSGDKVCEDEVLLIANHRTVDMMYFMDLIRKGOINITYVLKSSLMKL 133
QY 127 PLVGYFAOHGIIYVRSKAFNEKEMRNKLQSYVDAGTPLYVIFPEGTRYNDEQTVLS 186
DB 134 PLFGMAFHLFEFIPVERRWEDEANLRQIVSFPDPRDALMLALFPEGDTYTEAKCO--- 190
QY 187 ASQAFPAORGLAVLKHLVTRIKATHVAFDCMKNYLDATIDYVYVEGKDDGGORRESPT 246
DB 191 RSKRFAENGLPIILNNVLPRTKGFSVCLQELSCSDAVVDVITGYKTR----- 239
QY 247 MTEFLCKECP-----KIHIIHIDRDKDVEDEQEHMRMLHERFEIKDKMLIE 294
DB 240 -----CPSFLDNVYGIPESEVHHIRKINLTQIPNOEKDINAMLMNTFOLKDLND 291
QY 295 FYESPDERRRKRPKGSVNSKLSIKTLPMSLILSGLT-----AGMLMTDAGRKL 345
DB 292 FYSN-----GHPNDETEKEFNTKYLINCLAVIAFTTICTHLPFSSMIWF---RIY 341
QY 346 VNTWYIGTLIGCLMWT 361
DB 342 VS-----LACVYLT 350

RESULT 9
US-09-935-625-23025
; Sequence 23025, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 33025
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..375
; OTHER INFORMATION: Ceres Seq. ID no. 3443248
US-09-935-625-23025

Query Match          20.3%; Score 390.5; DB 5; Length 375;
Best Local Similarity 26.9%; Pred. No. 5.1e-28;
Matches 101; Conservative 78; Mismatches 128; Indels 69; Gaps 9;
```

```
DB 134 PLFGMAFHLFEFIPVERRWEDEANLRQIVSFPDPRDALMLALFPEGDTYTEAKCO--- 190
QY 187 ASQAFPAORGLAVLKHLVTRIKATHVAFDCMKNYLDATIDYVYVEGKDDGGORRESPT 246
DB 191 RSKRFAENGLPIILNNVLPRTKGFSVCLQELSCSDAVVDVITGYKTR----- 239
QY 247 MTEFLCKECP-----KIHIIHIDRDKDVEDEQEHMRMLHERFEIKDKMLIE 294
DB 240 -----CPSFLDNVYGIPESEVHHIRKINLTQIPNOEKDINAMLMNTFOLKDLND 291
QY 295 FYESPDERRRKRPKGSVNSKLSIKTLPMSLILSGLT-----AGMLMTDAGRKL 345
DB 292 FYSN-----GHPNDETEKEFNTKYLINCLAVIAFTTICTHLPFSSMIWF---RIY 341
QY 346 VNTWYIGTLIGCLMWT 361
DB 342 VS-----LACVYLT 350

RESULT 10
US-09-935-625-11062
; Sequence 11062, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 11062
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..351
; OTHER INFORMATION: Ceres Seq. ID no. 1394371
US-09-935-625-11062

Query Match          20.1%; Score 387.5; DB 5; Length 351;
Best Local Similarity 27.1%; Pred. No. 8.9e-28;
Matches 100; Conservative 75; Mismatches 125; Indels 69; Gaps 9;

QY 19 VLLGTAFTYVLAMG-----VWRLSAFLPAREYQALDRLCYQOSMVLFEFFENTGVQ 73
DB 1 MVLVSTAFMMLIFWGFLSAVVLRLFS---IRSRKCVSFFGSMWALMPFLFEKINKTK 56
QY 74 ILLYGLPRKKNIIYLANHOSVVDWIVADILAIRONALGHVRYVLEKGLKMLPLGYCYF 133
DB 57 VIFSGDKVPCEDRVLLIANHRTVDMMYFMDLIRKGOINITYVLKSSLMKLPLFSGMAF 116
QY 134 AQHGGIYVRSKAFNEKEMRNKLQSYVDAGTPLYVIFPEGTRYNDEQTVLSASQFAA 193
DB 117 HLEFIPVERRWEDEANLRQIVSFPDPRDALMLALFPEGDTYTEAKCO---RSKFAA 173
QY 194 QRGIAVLKHLVTRIKATHVAFDCMKNYLDATIDYVYVEGKDDGGORRESPTMTEFLCK 253
DB 174 ENGLPIILNNVLPRTKGFSVCLQELSCSDAVVDVITGYKTR----- 215
QY 254 ECP-----KIHIIHIDRDKDVEDEQEHMRMLHERFEIKDKMLIEFYESPDP 301
DB 216 -CPSFLDNVYGIPESEVHHIRKINLTQIPNOEKDINAMLMNTFOLKDLNDFYSN-- 271
QY 302 ERRKRPFGKSVNSKLSIKTLPMSLILSGLT-----AGMLMTDAGRKLVTWYIG 352
DB 272 -----GHPNDETEKEFNTKYLINCLAVIAFTTICTHLPFSSMIWF---RIYVS----- 319
QY 353 TLLGCLMWT 361
DB 320 --LACVYLT 326
```

```
RESULT 11
; Sequence 14198, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ. ID NOS: 33136
; SEQ ID NO 14198
; LENGTH: 351
; TYPE: PR1
; ORGANISM: Arabidopsis thaliana
; NAME/KEY: peptide
; LOCATION: 1..351
; OTHER INFORMATION: Ceres Seq. ID no. 3441315
US-09-935-625-14198
```

```
Query Match 20.1%; Score 387.5; DB 5; Length 351;
Best Local Similarity 27.1%; Pred. No. 8.9e-28;
Matches 100; Conservative 75; Mismatches 125; Indels 69; Gaps 9;
```

```
QY 19 VLLGTAPTYVLANG-----VWRLSAFLPARFYQALDDRLCYCYOSWVLFEEFNTGYQ 73
D 1 MVLVSTAFMMLIFWGFSAVVLRLFS---IRSRKCVSFFFGSWLWLPFLFEKINKTK 56
QY 74 ILLYGDLPKKNKENTIIYLANHOSVDWIVADILAIRONALGHVRYVLKGLKWLPLCYGYF 133
D 57 VIFSGDKVPCEDRVLLANHRTVEVDMWFMDLARKGQIGIKYVLSLKLPLFGWAF 116
QY 134 AQHGGIYVRSRANKNEKEMRKLOSVDAGTPMYLVIFPECTRYNPEQTKVLSASQAF 193
D 117 HLFEPFIVERMEVDEANLRQIVSFPDRDALMLALFPEGTDTYEAQCQ---RSKRFAA 173
QY 194 ORGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIVDVTVYVEGKDGQGRRESPTWTEFLCK 253
D 174 ENGLPLINNVLLPRTKGFVSCLOELSCSLDAVYDVITGYKTR----- 215
QY 254 ECP-----KIHIDRIDKKDVPEQEHMRRLHEREIKDKMLIEFYESPDP 301
D 216 -CPSEFLDNVGIPESEVHIHRIINLQIPNOEKDINAMLMNTFOLKQDLNDEYSN--- 271
QY 302 ERRRPFQKSVNSKLSIKTLPMSLLISGLT-----AGMLMTDAGRRLYVNTWYIG 352
D 272 ---GHFPEGTEKEFENRKTYLINCLAVIAFTTICTHLPFFSSMTWF---RIYVS----- 319
QY 353 TLLGCLMWT 361
D 320 --LACVYLT 326
```

```
RESULT 12
; Sequence 14202, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ. ID NOS: 33136
; SEQ ID NO 14202
; LENGTH: 351
; TYPE: PR1
; ORGANISM: Arabidopsis thaliana
; NAME/KEY: peptide
; LOCATION: 1..351
; OTHER INFORMATION: Ceres Seq. ID no. 3443249
US-09-935-625-14202
```

```
; NAME/KEY: peptide
; LOCATION: 1..351
; OTHER INFORMATION: Ceres Seq. ID no. 3441319
US-09-935-625-14202
```

```
Query Match 20.1%; Score 387.5; DB 5; Length 351;
Best Local Similarity 27.1%; Pred. No. 8.9e-28;
Matches 100; Conservative 75; Mismatches 125; Indels 69; Gaps 9;
```

```
QY 19 VLLGTAPTYVLANG-----VWRLSAFLPARFYQALDDRLCYCYOSWVLFEEFNTGYQ 73
D 1 MVLVSTAFMMLIFWGFSAVVLRLFS---IRSRKCVSFFFGSWLWLPFLFEKINKTK 56
QY 74 ILLYGDLPKKNKENTIIYLANHOSVDWIVADILAIRONALGHVRYVLKGLKWLPLCYGYF 133
D 57 VIFSGDKVPCEDRVLLANHRTVEVDMWFMDLARKGQIGIKYVLSLKLPLFGWAF 116
QY 134 AQHGGIYVRSRANKNEKEMRKLOSVDAGTPMYLVIFPECTRYNPEQTKVLSASQAF 193
D 117 HLFEPFIVERMEVDEANLRQIVSFPDRDALMLALFPEGTDTYEAQCQ---RSKRFAA 173
QY 194 ORGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIVDVTVYVEGKDGQGRRESPTWTEFLCK 253
D 174 ENGLPLINNVLLPRTKGFVSCLOELSCSLDAVYDVITGYKTR----- 215
QY 254 ECP-----KIHIDRIDKKDVPEQEHMRRLHEREIKDKMLIEFYESPDP 301
D 216 -CPSEFLDNVGIPESEVHIHRIINLQIPNOEKDINAMLMNTFOLKQDLNDEYSN--- 271
QY 302 ERRRPFQKSVNSKLSIKTLPMSLLISGLT-----AGMLMTDAGRRLYVNTWYIG 352
D 272 ---GHFPEGTEKEFENRKTYLINCLAVIAFTTICTHLPFFSSMTWF---RIYVS----- 319
QY 353 TLLGCLMWT 361
D 320 --LACVYLT 326
```

```
RESULT 13
; Sequence 14255, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ. ID NOS: 33136
; SEQ ID NO 14255
; LENGTH: 351
; TYPE: PR1
; ORGANISM: Arabidopsis thaliana
; NAME/KEY: peptide
; LOCATION: 1..351
; OTHER INFORMATION: Ceres Seq. ID no. 3443249
US-09-935-625-14255
```

```
Query Match 20.1%; Score 387.5; DB 5; Length 351;
Best Local Similarity 27.1%; Pred. No. 8.9e-28;
Matches 100; Conservative 75; Mismatches 125; Indels 69; Gaps 9;
```

```
QY 19 VLLGTAPTYVLANG-----VWRLSAFLPARFYQALDDRLCYCYOSWVLFEEFNTGYQ 73
D 1 MVLVSTAFMMLIFWGFSAVVLRLFS---IRSRKCVSFFFGSWLWLPFLFEKINKTK 56
QY 74 ILLYGDLPKKNKENTIIYLANHOSVDWIVADILAIRONALGHVRYVLKGLKWLPLCYGYF 133
D 57 VIFSGDKVPCEDRVLLANHRTVEVDMWFMDLARKGQIGIKYVLSLKLPLFGWAF 116
```





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:08:57 ; Search time 39.13 Seconds

(without alignments)  
893.854 Million cell updates/sec

Title: US-09-853-526-5  
Perfect score: 1924  
Sequence: 1 MLSTLVHTYSMRYLPSV.....YVNTWIVGTLLGLMTWITKA 364

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	470	24.4	918	2	T34057	hypothetical prote
2	398.5	20.7	393	2	B96780	hypothetical prote
3	328	17.0	374	2	S52645	probable 1-acyl-gl
4	307	16.0	376	2	S60478	probable 1-acyl-sn
5	263	13.7	376	2	D96550	hypothetical prote
6	249.5	13.0	350	2	T40466	probable acetyltra
7	242.5	12.6	344	2	T31913	hypothetical prote
8	239	12.4	311	2	T07936	probable glycerol-
9	238.5	12.4	310	2	T06755	probable glycerol-
10	231.5	12.0	397	2	S45900	probable membrane
11	221.5	11.5	396	2	S54641	probable membrane
12	213.5	11.1	523	2	T25998	hypothetical prote
13	212	11.0	439	2	T22689	hypothetical prote
14	184	9.6	295	2	B83587	hypothetical prote
15	182	9.5	391	2	T15366	hypothetical prote
16	164.5	8.5	310	2	S40808	polynucleotide ade
17	164.5	8.5	310	2	A91227	probable endonucle
18	164.5	8.5	310	2	G86073	probable endonucle
19	153	8.0	304	2	B83541	probable polynucle
20	145.5	7.6	285	2	T15252	hypothetical prote
21	136.5	7.1	302	2	A10950	probable acyltrans
22	131.5	6.8	363	2	T20608	hypothetical prote
23	130	6.8	294	2	D82371	probable polynucle
24	124	6.4	303	2	A48600	probable sn2-acylg
25	118	6.1	281	2	S60477	1-acylglycerol-3-p
26	116	6.0	266	2	S73863	probable 1-acylgly
27	114.5	6.0	247	2	G72223	hypothetical prote
28	112	5.8	258	2	G83101	probable acyltrans
29	111.5	5.8	377	1	A44216	major envelope ant

30	111.5	5.8	488	1	QOEBHS	alkaline exonuclea
31	109.5	5.7	488	2	T44030	alkaline exonuclea
32	109.5	5.7	488	2	T44215	alkaline exonuclea
33	109	5.7	240	2	D64089	1-acylglycerol-3-p
34	108	5.6	261	2	C82067	probable 1-acylgly
35	104	5.4	240	2	D64688	probable 1-acylgly
36	104	5.4	267	2	AC3499	1-acyl-sn-glycerol
37	103	5.4	243	2	AC0083	1-acylglycerol-3-p
38	103	5.4	266	2	B98283	hypothetical prote
39	103	5.4	266	2	AG3000	1-acyl-sn-glycerol
40	102	5.3	212	2	G85357	hypothetical prote
41	102	5.3	1551	2	T18941	hypothetical prote
42	101	5.2	262	2	T22599	1-acylglycerol-3-p
43	101	5.2	323	2	C83940	sugar ABC transport
44	99.5	5.1	237	2	B71827	probable 1-acylgly
45	99	5.1	225	2	B45582	probable 1-acylgly

ALIGNMENTS

RESULT	1	
T34057		hypothetical protein F28B3.5 - Caenorhabditis elegans
C:Species:	Caenorhabditis elegans	
C:Date:	29-Oct-1999	#sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession:	T34057	
R:Geisel, C.; Kramer, J.; Smith, A.		
A:Submitted to the EMBL Data Library, May 1997		
A:Description:	The sequence of C. elegans cosmid F28B3.	
A:Reference number:	221469	
A:Accession:	T34057	
A:Status:	preliminary; translated from GB/EMBL/DBJ	
A:Molecule type:	DNA	
A:Residues:	1-918 <GEI>	
A:Cross-references:	EMBL:AF003136; PIDN:AAB93636.1; GSPDB:GN00019; CESP:F28B3.5	
A:Experimental source:	strain Bristol N2; clone F28B3	
C:Genetics:		
A:Gene:	CESP:F28B3.5	
A:Map position:	1	
A:Introns:	85/3; 129/3; 235/3; 418/2; 482/3; 532/3; 736/3; 829/3; 850/2	
Query Match	24.4%	Score 470; DB 2; Length 918;
Best Local Similarity	33.0%	Pred. No. 5.8e-32;
Matches 131; Conservative 75; Mismatches 143; Indels 48; Gaps 14;		
QY	1	MLSTLVHTYSMRYLIP---SVLLGTAPTYVLAWGVMRLSAFLPAREYQALDDRLYC 56
DB	533	MLSTLVHTYSMRYLIP---SVLLGTAPTYVLAWGVMRLSAFLPAREYQALDDRLYC 56
QY	57	VYGSNVLFFRENTGTVOILLXGDLK--NK-----ENIIYLANHOSYVDWIVADILAIRON 110
DB	589	SYMLCLFFVENISGVEIYLGHNIEVNTGKRENNVMSINQSNVDIIPVLAARHG 648
QY	111	ALGH---VRVVLKGLKWLPLCYGCPFAOHCIIYVRSKAKNEKMRKLSYVDAGTPMY 167
DB	649	DQGNQAFRVWVKNSIHLVMEFGYITFOHGITYRRGEGITGAPVLRQLKMLNESDPYW 708
QY	168	LVTFEGSTRYNPBOTKVASAQAFAOGLAVLKRVLPRIKATHVAFDCKMYLDAIYD 227
DB	709	LVTFEGSTRYNPBOTKVASAQAFAOGLAVLKRVLPRIKATHVAFDCKMYLDAIYD 227
QY	228	VYVAVY-EGKDDG-----QRESPTWFLC--KCPKIHIIHIDIK 267
DB	768	VYVAVY-EGKDDG-----QRESPTWFLC--KCPKIHIIHIDIK 267
QY	268	KDVPBEGEHMRMLHEFEIKDKMLIEFYSPDERKRRPKGSVNSKLSIKKTLPSMLI 327
DB	828	DEVPKARLELRTWTTEFTFKERIDIFY-SEKPSGSGALP-----CVPISQTLPSLTF 880
QY	328	LS-GTFAGMLMTDAGRKLYVNTWIVGTLLGLCLMTWITK 363

Db 881 FSAALLAPFESRTTGR-IYLLTITASSPLL-IAMLHIR 915

RESULT 2

B96780  
hypothetical protein F9E10.13 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: B96780  
R:Theologias, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso-  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Cressy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizart, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Traiser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719  
A:Accession: B96780  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-393 <STO>  
A:Cross-references: GB:AE005173; MID:96646762; PIDN:AAF21074.1; SSPDB:GN00141  
C:Genetics:  
A:Gene: F9E10.13  
A:Map position: 1

Query Match 20.7%; Score 398.5; DB 2; Length 393;  
Best Local Similarity 29.8%; Pred. No. 2.5e-26;  
Matches 97; Conservative 68; Mismatches 117; Indels 43; Gaps 7;

12 MRYLLPSVVLGTAPTYLANGVWRLLSAFLPARFYQALDRLCYQVSNVL----- 63  
Db LRGMLILVFLSTAPMFL-----YFAP---IALGLRLSLVQDSKRVSLIFGLW 70  
QY 64 -----FFENTYGVQILYGLDLPKKNENIILANHOSTVDWIVADILAIROMALGHVRYV 118  
Db LALMPYLEFTYNGTGVVSSGDIIPVEKRVLLIANHREVDWMLNIALRGCGGIIYV 130  
QY 119 LKEGLKMLPLGYCYFAOHGCIYKRSKAFENKEMRNKQSYVDAGTPMYLVIPFEGTKYN 178  
Db LKSSLMKLPFGWGFHYLDEFIPVERKREVDPLQLMLSPKDOEPMLALFPEGIDFT 190  
QY 179 PEQTKVLASQAFAPAOAGLAVLKHVLPRIKATHVAFPCMKNYLDAYDVVVYEGKDDG 238  
Db 191 EEKCK---RSQKFAVEGLPALSNVLLPKTKGFGVCLVLSIDAVDLTIATKPR-- 244  
QY 239 GQRESPTMTFELCKECP-KIHIIHDIRIDKDVPEDEQEHMRWLHERFEIKDKMLIEFY 297  
Db 245 -----CSPFMNVFGTDPSEVHIHVRRLKLEIFANEAESSAMLMDSKFLDKLSDF-- 297  
QY 298 SPDEPRKRRFPCKSVNSKLITKTL 322  
Db 298 ----NAQCKFPNQRPEELSVLKCI 318

RESULT 3

S52645  
probable 1-acyl-glycerol-3-phosphate acyltransferase - maize  
C:Species: Zea mays (maize)  
C>Date: 14-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 19-May-2000  
C:Accession: S52645  
R:BROWN, A.P.; Coleman, J.; Tomney, A.M.; Watson, M.D.; Slabas, A.R.  
Plant Mol. Biol. 26, 211-223, 1994  
A:Title: Isolation and characterization of a maize cDNA that complements a 1-acyl sn-gly-  
cer acyltransferases.  
A:Reference number: S52645; MUID:95035993  
A:Accession: S52645  
A:Status: preliminary  
A:Molecule type: mRNA

A:Residues: 1-374 <BRO>  
A:Cross-references: EMBL:Z29518; MID:9575959; PIDN:CAA82638.1; PID:9575960  
C:Superfamily: probable membrane protein YBR042c

Query Match 17.0%; Score 328; DB 2; Length 374;  
Best Local Similarity 29.8%; Pred. No. 2.7e-20;  
Matches 92; Conservative 64; Mismatches 121; Indels 32; Gaps 8;

29 VLAMGWRLLSA-----FLPAR-----FYQALDRLCYQVSNVLFPEFNTYGVQ 73  
Db VLPGLGLEFLSLGLVNAIQAVLFYIRPFSKSTIRIRFLAELIMQLVAVVDMAGVK 68  
QY 74 ILLYGD----LPKKNENIILANHOSTVDWIVADILAIROMALGHVRYLKEGLKMLPLY 129  
Db VOLAHDEFTYSMCKEHLIISNRHSDIDWILGIMLAQRSCSLGTLAVMKSSKFLPYI 128  
QY 130 G--CYFAOHGCIYKRSKAFENKEMRNKQSYVDAGTPMYLVIPFEGTKYNPEQTKVLSA 187  
Db GWSMWFAYE--LFERSWAKDEKTLKQRLKDFPRPFMLALFEGTRFPAPK---LLA 183  
QY 188 SOAFAPQGLAVLKHVLPRIKATHVAFPCMKNYLDAYDVVVYEGKDDGQRESPTM 247  
Db 184 AOETASQGLPAPRNVLPRTKGFVSASIMKDEVPALYDTYVP-KDS-----PQPTM 237  
QY 248 TEFCKECPKIHIIHDIRIDKDVPEDEQEHMRWLHERFEIKDKMLIEFESPDERRKRF 307  
Db LRLKGSSVYIHWKMKRIANSEMPKSDSDVSKWCKDIFVANDALLDKLATGTDEELRP 297  
QY 308 PGKSVNSKL 316  
Db 298 IGRPVKSL 306

RESULT 4

S60478  
probable 1-acyl-sn-glycerol-3-phosphate acyltransferase - Limnanthes douglasii  
C:Species: Limnanthes douglasii  
C>Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 17-Mar-2000  
C:Accession: S60478  
R:BROWN, A.P.; Brough, C.L.; Kroon, J.T.M.; Slabas, A.R.  
Plant Mol. Biol. 29, 267-278, 1995  
A:Title: Identification of a cDNA that encodes a 1-acyl-sn-glycerol-3-phosphate acyl-  
A:Reference number: S60477; MUID:96046746  
A:Accession: S60478  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-377 <BRO>  
A:Cross-references: EMBL:Z48730; MID:91067137; PIDN:CAA88620.1; PID:91067138  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1995  
C:Superfamily: probable membrane protein YBR042c

Query Match 16.0%; Score 307; DB 2; Length 377;  
Best Local Similarity 25.5%; Pred. No. 1.7e-18;  
Matches 91; Conservative 85; Mismatches 131; Indels 50; Gaps 13;

37 LLSAFLPARFY---QALDRLCYQVSNV-----LFFENTYGVQILYGLDLPK--- 82  
Db LVNFIQAVFYVLPRIKATHVAFPCMKNYLDAYDVVVYEGKDDGQRESPTMTFELCKECPKI 80  
QY 83 -NKENIILANHOSTVDWIVADILAIROMALGHVRYLKEGLKMLPLYG--CYFAOHGCI 139  
Db 81 MGKEHALDICNHRSDIDWILGIMVLAQRSCSLSSIAVKKSSKFLPYIGSMFSEY--L 138  
QY 140 VYKSAKNEKEMRNKQSYVDAGTPMYLVIPFEGTKYNPEQTKVLSA 198  
Db 139 FLERWAKDEKTLKQRLNDFPRFVLAFLVEGTRP---TRAKLLAAQETASADLP 194  
QY 199 VLKHVLPRIKATHVAFPCMKNYLDAYDVVVYEGKDDGQRESPTMTFELCKECPKI 258  
Db 195 VPRNVLPRIKGFVSASVNSMRSFPAIYDTYVAP-----KTEQPTMLALFRKSSVY 248













**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:17:22 ; Search time 21.99 Seconds

(without alignments)  
640.924 Million cell updates/sec

Title: US-09-853-526-5  
1924

Sequence: 1 MLSTLVHTSMRRLPSVY.....YVNWIVGTILGCLMTVTKA 364

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1871	97.2	353	1	P1CE_HUMAN
2	1568.5	81.5	354	1	P1CE_MOUSE
3	270	14.0	378	1	P1CD_RAT
4	268.5	14.0	378	1	P1CD_HUMAN
5	240	12.5	376	1	P1CC_HUMAN
6	231.5	12.0	397	1	P1B2_YEAST
7	221.5	11.5	396	1	YD18_YEAST
8	204.5	10.6	370	1	Y205_HUMAN
9	182	9.5	391	1	Y773_CAEL
10	164.5	8.5	310	1	Y1NG_ECOLI
11	124	6.4	303	1	P1SC_YEAST
12	119	6.2	281	1	P1SC_LIMDO
13	116	6.0	266	1	P1SC_MYCPN
14	116	6.0	281	1	P1SC_LIMAL
15	111.5	5.8	377	1	VENV_FOPVY
16	111.5	5.8	488	1	EXON_HSV62
17	109.5	5.7	488	1	EXON_HSV62
18	109	5.7	240	1	P1SC_HAREIN
19	104	5.4	240	1	P1SC_HELPI
20	104	5.4	308	1	P1SC_COCNU
21	102	5.3	1551	1	Y012_CAEL
22	101	5.2	262	1	P1C1_CAEL
23	99.5	5.2	237	1	P1SC_HELPI
24	99	5.1	245	1	P1SC_SALTY
25	98	5.1	484	1	Y142_NPVO
26	97	5.0	245	1	P1SC_ECOLI
27	94.5	4.9	282	1	P1C2_CAEL
28	93	4.8	280	1	P1SC_BORBU
29	90.5	4.7	1241	1	TRK1_SACBA
30	90	4.7	892	1	RA16_SCHPO
31	88.5	4.6	446	1	M164_MOUSE
32	87.5	4.5	503	1	CPN1_PAPHA
33	86.5	4.5	268	1	P1SC_MYCCE

34	86	4.5	493	1	SYK_ARCFU
35	85.5	4.4	160	1	ALGO_PSEAE
36	85	4.4	561	1	5NTC_HUMAN
37	84.5	4.4	255	1	P1SC_NEIMB
38	84.5	4.4	383	1	P46E_ARATH
39	84	4.4	1272	1	Y228_METJA
40	83.5	4.3	255	1	P1SC_NEIGO
41	83.5	4.3	255	1	P1SC_NEIMA
42	83.5	4.3	503	1	CPN1_HUMAN
43	83.5	4.3	2386	1	RAD3_SCHPO
44	83	4.3	318	1	YOK3_CAEL
45	83	4.3	778	1	YF05_METJA

## ALIGNMENTS

RESULT	ID	PLCE_HUMAN	STANDARD	PRT	353 AA.
AC	09NU02	09B0G4			
DT	01-MAR-2002	(Rel. 41, Created)			
DT	01-MAR-2002	(Rel. 41, Last sequence update)			
DT	01-MAR-2002	(Rel. 41, Last annotation update)			
DE	1-acyl-sn-glycerol-3-phosphate acyltransferase epsilon (EC 2.3.1.51)				
DE	(1-AGP acyltransferase 5) (1-AGPAT 5) (Lysophosphatidic acid acyltransferase-epsilon) (LPAAT-epsilon) (1-acylglycerol-3-phosphate O-acyltransferase 5).				
DE	AGPAT5.				
GN	Homo sapiens (Human).				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Leung D.W.;				
RT	"Cloning and expression of LPAT-epsilon.";				
RL	Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RA	Cohen D., Chumakov I., Blumenfeld M., Bougueleret L.;				
RL	Patent number WO9332644, 01-JUL-1999.				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Amalgam;				
RX	MEDLINE=21154917; PubMed=11230166;				
RA	Wiemann S., Well B., Wellenreuther R., Gassenhuber J., Glassl S.,				
RA	Ansorge W., Boecker M., Bloeker H., Bauersachs S., Blum H.,				
RA	Lauber J., Duesterhoeft A., Beyer A., Koehler K., Strack N.,				
RA	Mewes H.-W., Othenwaelter B., Obermaier B., Tampe J., Heubner D.,				
RA	Wambutt R., Korn B., Klein M., Poustka A.;				
RT	"Towards a catalog of human genes and proteins: sequencing and				
RL	analysis of 500 novel complete protein coding human cDNAs.";				
RL	Genome Res. 11:422-435(2001).				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Placenta.				
RA	Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,				
RA	Nishikawa T., Nagai R., Sugano S., Shiratori A., Sudo H.,				
RA	Wagatsuna M., Hosoliri T., Raku Y., Kodaira H., Kondo H., Sugawara M.,				
RA	Takahashi M., Chiba Y., Ishida S., Kodakawa K., Ono Y., Takiguchi S.,				
RA	Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,				
RA	Yamanoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y.,				
RA	Nimomiya K., Iwayanagi T.;				
RT	"NEO human cDNA sequencing project.";				
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.				
CC	- FUNCTION: CONVERTS LYSOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC				
CC	ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION (BY				
CC	SIMILARITY).				
CC	- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate -				
CC	CoA + 1,2-diacyl-sn-glycerol 3-phosphate.				
CC	- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.				
CC	- SUBCELLULAR LOCATION: Integral membrane protein (potential).				

```

CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE
CC ACYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF375789; AKS4809.1; ALT_INT.
DR EMBL: AL136587; CAB66322.1; ALT_INT.
DR EMBL: AK002072; BAA92069.1; -.
DR InterPro: IPR002123; Acyltransferase.
DR Pfam: PF01553; Acyltransferase; 1.
KM Phospholipid biosynthesis; Transferase; Acyltransferase;
KM Transmembrane.
FT TRANSMEM 7 29 POTENTIAL.
FT TRANSMEM 44 66 POTENTIAL.
FT TRANSMEM 334 351 POTENTIAL.
FT CONFLICT 145 145 L -> V (IN REF. 2).
SQ SEQUENCE 353 AA; 40813 MW; A05B1FA246CE1B64 CRC64;

Query Match 97.2%; Score 1871; DB 1; Length 353;
Best Local Similarity 100.0%; Pred. No. 1.3e-151;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 MRYLLPSVLLGTAPTYVLAWGVRLLSAFLPARFYQALDRLCYQVOSVLEFFENYTG 71
DB 1 MRYLLPSVLLGTAPTYVLAWGVRLLSAFLPARFYQALDRLCYQVOSVLEFFENYTG 60
QY 72 VOILLYGDLPKKNENIYYLANHSTVDMIVADILAIQNALGHRVYLKESGKMLP.YXG 131
DB 61 VOILLYGDLPKKNENIYYLANHSTVDMIVADILAIQNALGHRVYLKESGKMLP.YXG 120
QY 132 YFAOHGCIYVRSKAKFENKEKRNKLSQSYVDAGTDMYLVIFPEGTRVNPEDTKVLSAQAF 191
DB 121 YFAOHGCIYVRSKAKFENKEKRNKLSQSYVDAGTDMYLVIFPEGTRVNPEDTKVLSAQAF 180
QY 192 AAQGLAVLKHVLPRIKATKVAFDCKKNYLDATYDVTYVEGKDDGQRRSPPTWTEFL 251
DB 181 AAQGLAVLKHVLPRIKATKVAFDCKKNYLDATYDVTYVEGKDDGQRRSPPTWTEFL 240
QY 252 CKECPKIHIDRIKDKVDVEQEHMRRLHEREFEIKDKMLIEFYESPDPERRKRPFGKS 311
DB 241 CKECPKIHIDRIKDKVDVEQEHMRRLHEREFEIKDKMLIEFYESPDPERRKRPFGKS 300
QY 312 VNSKLSIKTKLPMSLLISGLTAGMLMTDAGRKLYVNTWYIGTLGCLWVITKA 364
DB 301 VNSKLSIKTKLPMSLLISGLTAGMLMTDAGRKLYVNTWYIGTLGCLWVITKA 353

RESULT 2
PLOC_MOUSE STANDARD; PRT; 354 AA.
AC 09D1E8;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DE 1-acyl-sn-glycerol-3-phosphate acyltransferase epsilon (EC 2.3.1.51)
DE (1-acyl acyltransferase 5) (1-AGPAT 5) (lysophosphatidic acid
DE acyltransferase-epsilon) (1-AGPAT-epsilon) (1-acylglycerol-3-phosphate
DE O-acyltransferase 5).
GN AGPAT5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;

```

```

RX MEDLINE=21085660; Pubmed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Atzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamamaki I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Kleihsman W., Gaasterland T., Glasi C., King B., Kochua H.,
RA Kiehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schirral L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Mashio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barch G.,
RA Blake J., Boffelli D., Bojunga N., Carrinci P., de Bonaio M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmig L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -1- FUNCTION: CONVERTS LYSOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC
CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION (BY
CC SIMILARITY).
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate -
CC COA + 1,2-diacyl-sn-glycerol 3-phosphate.
CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE
CC ACYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AK003649; BAB2915.1; -.
DR MGD: MGI:1915880; Agpat5.
DR InterPro: IPR002123; Acyltransferase.
DR Pfam: PF01553; Acyltransferase; 1.
KM Phospholipid biosynthesis; Transferase; Acyltransferase;
KM Transmembrane.
FT TRANSMEM 7 29 POTENTIAL.
FT TRANSMEM 44 66 POTENTIAL.
FT TRANSMEM 335 352 POTENTIAL.
SQ SEQUENCE 354 AA; 40943 MW; 447EBF9249B1E800 CRC64;

Query Match 81.5%; Score 1568.5; DB 1; Length 354;
Best Local Similarity 81.4%; Pred. No. 6.5e-126;
Matches 288; Conservative 39; Mismatches 26; Indels 1; Gaps 1;

QY 12 MRYLLPSVLLGTAPTYVLAWGVRLLSAFLPARFYQALDRLCYQVOSVLEFFENYTG 71
DB 1 MRYLLPSVLLGTAPTYVLAWGVRLLSAFLPARFYQALDRLCYQVOSVLEFFENYTG 60
QY 72 VOILLYGDLPKKNENIYYLANHSTVDMIVADILAIQNALGHRVYLKESGKMLP.YXG 131
DB 61 VOILLYGDLPKKNENIYYLANHSTVDMIVADILAIQNALGHRVYLKESGKMLP.YXG 120
QY 132 YFAOHGCIYVRSKAKFENKEKRNKLSQSYVDAGTDMYLVIFPEGTRVNPEDTKVLSAQAF 191
DB 121 YFAOHGCIYVRSKAKFENKEKRNKLSQSYVDAGTDMYLVIFPEGTRVNPEDTKVLSAQAF 180
QY 192 AAQGLAVLKHVLPRIKATKVAFDCKKNYLDATYDVTYVEGKDDG-QGRRSPPTWTEFL 250
DB 181 AAQGLAVLKHVLPRIKATKVAFDCKKNYLDATYDVTYVEGKDDG-QGRRSPPTWTEFL 240
QY 251 CKECPKIHIDRIKDKVDVEQEHMRRLHEREFEIKDKMLIEFYESPDPERRKRPFGK 310
DB 241 CKECPKIHIDRIKDKVDVEQEHMRRLHEREFEIKDKMLIEFYESPDPERRKRPFGK 300

```





```

Db 133 GMMWFTM--VFCEKMEODKRYATSLQHLRDVPEKYPFLIHCEGRTEKKEI---187
Oy 188 SOAFAGGLAVLKHVLPRIKATHVADFCMKNDYADTVVYEGRDGGRRESPTM 247
Db 188 SMOVARAKGLRKLHLLPRKGAFTYRSLSRNVSAVYDCLTNF-----FNNEFTL 240
Oy 248 TELCKECPKHIIHIDRIKDKVPEEGRMRMRHREFEIKMLIEFY-----ESP-D 300
Db 241 LGVLNGKRYHADLVYRIRPLEDPEDEDCSAMLHKLYOEKDAFQEEYRTGTFETPMV 300
Oy 301 PRRKRRKRSVN 313
Db 301 PPR--PWTLVN 310

RESULT 5
PLCC_HUMAN STANDARD; PRT; 376 AA.
AC 09NR27; 09NR26;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE 1-acyl-sn-glycerol-3-phosphate acyltransferase gamma (EC 2.3.1.51) (1-
DE ACP acyltransferase 3) (1-ACP 3) (lysophosphatidic acid
DE acyltransferase-gamma) (LPAAT-gamma) (1-acylglycerol-3-phosphate O-
GN ACPAT3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS GAMMA-1 AND GAMMA-2).
RA Leung D.W.;
RT "Structure and functions of lysophosphatidic acid acyltransferases.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA TISSUE-Fetal liver;
RA Nagamine K., Kudoh J., Minoshima S., Kawasaki K., Hase T.,
RA Shimizu N.;
RT "Isolation of a novel gene encoding 1-acylglycerol-3-phosphate
RT O-acyltransferase 3 (ACPAT3) from the human chromosome 21q22.3.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA TISSUE-Brain;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CONVERTS LYOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC
CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION (BY
CC SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ACPYL-COA + 1-acyl-sn-glycerol 3-phosphate =
CC COA + 1,2-diacyl-sn-glycerol 3-phosphate.
CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: GAMMA-1 (SHOWN HERE) AND GAMMA-
CC 2; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE
CC ACYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF156774; AAF80336.1; -
DR EMBL; AF156775; AAF80337.1; -
DR EMBL; AB040138; BAB18943.1; -

```

```

DR EMBL; BC011971; AAH11971.1; -
DR InterPro; IPR002123; Acyltransferase.
DR Pfam; PF01553; Acyltransferase; 1.
KW Phospholipid biosynthesis; Transferase; Acyltransferase;
KW Transmembrane; Alternative splicing.
FT TRANSMEM 15 39 POTENTIAL.
FT TRANSMEM 126 144 POTENTIAL.
FT TRANSMEM 308 329 POTENTIAL.
FT TRANSMEM 336 352 POTENTIAL.
FT VARSPLIC 1 62 MISSING (IN ISOFORM GAMMA-2).
SQ SEQUENCE 376 AA; 43381 MW; C12CDBB/C363852 CRC64;

Query Match 12.5%; Score 240; DB 1; Length 376;
Best Local Similarity 25.4%; Pred. No. 5.5e-13;
Matches 69; Conservative 63; Mismatches 116; Indels 24; Caps 7;

Oy 45 RFQALDRLYCVQSMVLPFFPENTGVQIIIXDLPK-----NKENITYLANHSTYDVI 100
Db 44 QLYRLNCRALVSLWSQVLMLEWMSCTECTLFTDOATVERFGREHAVIILNHNFEIDL 103
Oy 101 VADILAIROMLAGHYRVLYKREGKMLPLYG--CYFAQHGIIYVRSKAFNEKMRNKLOS 158
Db 104 CGWTKCEPFGVLSKVLAKKELVPLIGTWYFLEI--VFCKRKEEDDTYVEGLR 161
Oy 159 YVDAGTPMYLVIFPEGTRYNEQTRKVISASQAFAAQRGLAVLKHVLPRIKATHVADCM 218
Db 162 LSDPEYWMFLLYCEGTFTETKRV--SMEVAAKGLPVLYKHYLLPRTGFTTAVACL 218
Oy 219 KNYDAIYDVVYVYEGKBDGGRRESPTMTFEFLCKECKKHIIHIDRIKDKVPEEGRMR 278
Db 219 KGYVAAYVDVTLNFRG-----NNPSSLGILTKKYEADMVRRPLEDIPDEKAA 271
Oy 279 RWLHREFEIKMKLIEFESPPERRKRPCK 310
Db 272 QMLHKLVEKD-ALQEIY-----NQKGMFPGE 297

RESULT 6
IDB2_YEAST STANDARD; PRT; 397 AA.
AC P38226;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 45.5 kDa protein in FAY1-TCM62 intergenic region.
DE YBR042C OR YBR0412.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-S288C;
RA Andre B., Cziepluch C., Hein C., Jauniaux J.-C., Urrestarazu A.,
RA Vissers S.;
RT Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RL -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE
CC ACYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z35911; CA84984.1; -
DR PIR; S45900; S45900.
DR SGD; S0000246; YBR042C.
DR InterPro; IPR002123; Acyltransferase.

```

DR	Pfam; PF01553; Acyltransferase; 1.
KW	Hypothetical protein; Phospholipid biosynthesis; Transferase;
FT	Acyltransferase; Transmembrane.
FT	TRANSMEM 13 33 POTENTIAL.
FT	TRANSMEM 58 78 POTENTIAL.
FT	TRANSMEM 118 138 POTENTIAL.
FT	TRANSMEM 377 397 POTENTIAL.
SO	SEQUENCE 397 AA; 45515 MW; 04CADD1247D98A8 CMC64;

Query Match	12.0%;	Score 231.5;	DB 1;	Length 397;
Best Local Similarity	25.9%;	Pred. No. 3.1e-12;		
Matches	88;	Conservative	55;	Mismatches 136;
			Indels	61;
			Gaps	13;

QY	54	LYCYOVSUVELEFFENYNGVOILLGGDLPEKN-----KENIITYLANHOSIVMADIIAI	107
Db	67	LHVAPSAVRIITTESSVPKGTFFELDKKKRILSHLKSNSVAICHQIYTTDITWIFLMIAY	126
QY	108	RONALGHVRYVLKGLKMLPLGYCYFAQHGIGYVRSKAFNEKERMKL-----	156
Db	127	TSNLGANWFILKKSLSASIPILYCGMRNYNFTFMSRKMAOKDITLSNLSLAGDSNARGAG	186
QY	157	-----OSYVDAGT-----PMYLVIPPEGTRYNEQOKVLSASAFMAQ	194
Db	187	SLAGKSPERTEEGBSTINPEYIDPKQIHMPIYNNLLPEEGTNLSAD-TRQKSAK--YAAK	243
QY	195	RGLAVLKHVLPRIKATHVAEDCKNNYLDAIYDVIVVEG---KDDGGORRESPTMEFLC	252
Db	244	IGKKPKFKNVLLPHSGTLRYSLQKLPKSIESLYDITIGYSGVQEYEGELIYG-LKSIPE	302
QY	253	KECKR-IHIIHIDRLDKVRP-EBOEHMRMRWHERETIDKMLIEYES---PDERKKR	306
Db	303	GKYRLVDIHRADYVDIPLEDERSEMLKTIKSEKDALMERIYSTGSEYSDPE---	358
QY	307	EPGKSVNSKISIKK-TIPSMILISGLTAGMLMTDAGRKY	345
Db	359	-TNHSVTDSEFKINRIELEVLITPLTITWY---YKLY	393

RESULT	7
ID18_YEAST	
ID18_YEAST	STANDARD; PRT; 396 AA.
AC	Q12185;
DT	15-JUL-1999 (Rel. 38, Created)
DT	15-JUL-1999 (Rel. 38, Last sequence update)
DT	15-JUL-1999 (Rel. 38, Last annotation update)
DE	Hypothetical 45.9 kDa protein in KCS1-GCVI intergenic region.
GN	YDR018C OR YD9335.04C OR PZF396.
OS	Saccharomyces cerevisiae (Baker's yeast).
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC	Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX	NCBI_TaxID=4932;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=S288C / AB972;
RA	Deidman K., Brown D., Hamlyn N., Bowman S., Barrell B.G.,
RA	Rajandream M.A.;
RL	Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
RN	[2]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=97051598; PubMed=8896275;
RT	Eide L.G., Sander C., Prydz H.;
RT	"Sequencing and analysis of a 35.4 kb region on the right arm of
RT	chromosome IV from Saccharomyces cerevisiae reveal 23 open reading
RT	frames. ";
RL	Yeast 12:11085-1090(1996).
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC	-1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE
CC	ACYLRANSEASE FAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed, usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb.sib.ch](mailto:license@isb.sib.ch)).

---

CC  
DR EMBL; Z74314; CAA98838.1; -  
DR EMBL; X95966; CAA65210.1; -  
DR EMBL; Z49770; CAA89843.1; -  
DR SGD; S0002425; YDR018C.  
DR InterPro: IPR002123; Acyltransferase.  
DR Pfam: PF01553; Acyltransferase; 1.  
KW Hypothetical protein, Phospholipid biosynthesis; Transferase;  
KW Acyltransferase; Transmembrane.  
FT TRANSMEM 27 47 POTENTIAL.  
FT TRANSMEM 69 89 POTENTIAL.  
FT TRANSMEM 123 143 POTENTIAL.  
FT TRANSMEM 392 392 POTENTIAL.  
SQ SEQUENCE 396 AA; 45938 MW; 9F55AB56C8FD44DD CRC64;

Query Match	11.5%;	Score 221.5;	DB 1;	Length 396;
Best Local Similarity	25.4%;	Pred. No. 2.2e-11;		
Matches 78; Conservative	49;	Mismatches 113;	Indels 67;	Gaps 10;

```

QY 84 KENTIIYLANHOSVDWIMYADIIATIQNALGNHRYVLYKEGKMLPLXGCFQA0HGIGYVR 143
Db 108 KDRAIIIIANHHQYADWIMYIYLMWLSFVSNLGGVNYIILKALQIYIPLLGEMRNFFKFLFLSR 167
QY 144 SAKFENEKEMRNKL-----QSYVDAGTMY-LVIFPEGSTRNPEQYK 183
Db 168 MWOKERKALTNLSVSMIDNARCQGLTVNKCSCYKTNESIAAYNLIIMPEGNLS--LK 224
QY 184 VLSASQAPAAQRL--ALKHVLPFRKATNHPFCOMKNYIDAIYDVMYVYSGKDDGGR 241
Db 225 TREKSEACQRAHNDHVDRLHLLPHSGKGLFAVKLAPSDAIYDVIIGY----- 275
QY 242 RESPTM-----TEFLKEC-----KRIIHIDRIDKKDVP-EEQEHMRMLHEREFI 287
Db 276 --SPALFREVYGTKEFTLKIFLPMGVYPEKVDREYREFRVNEIPLODDEFEFMMLLGVME 333
QY 288 KDKMLIEYESPDPERRRKRPFGKSNLSKTKT-----LPSWILSLGTAGML 336
Db 334 KDQLEDDIYNT-----GQFSNANKNDOSIYVTTQTTGFQHEHTLPRILSLTYGFAFALI 387
QY 337 MTDAGRK 343
Db 388 LVFVWK 394

```

RESULT	8
Y205_HUMAN	
ID_Y205_HUMAN	STANDARD; PRT; 370 AA.
AC	Q92604
DT	15-JUL-1999 (Rel. 38, Created)
DT	15-JUL-1999 (Rel. 38, Last sequence update)
DT	16-OCT-2001 (Rel. 40, Last annotation update)
DE	Hypothetical protein KIAA0205.
GN	KIAA0205.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Bone marrow;
RX	MEDLINE=71191544; PubMed=9039502;
RA	Nagase T., Seki N., Ishikawa K.-I., Ohira M., Kawabayashi Y.,
RA	Ohara O., Tanaka A., Kotani H., Miyajima N., Nomura N.;
RT	"Prediction of the coding sequences of unidentified human genes. VI.
RT	The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by
RT	analysis of cDNA clones from cell line KG-1 and Brain."
RL	DNA Res. 3:321-329(1996).
CC	-1 SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
CC ACYLTRANSFERASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: D86960; BAA13196.1; -  
DR InterPro: IPR002123; Acyltransferase.  
KW Hypothetical protein; Phospholipid biosynthesis; Transferase;  
KW Acyltransferase; Transmembrane.  
FT TRANSMEM 22 42 POTENTIAL.  
FT TRANSMEM 342 362  
SQ SEQUENCE 370 AA; 43089 MW; 93498544EA651541 CRC64;

Query Match 10.6%; Score 204.5; DB 1; Length 370;  
Best Local Similarity 22.8%; Pred. No. 5,6e-10;  
Matches 90; Conservative 66; Mismatches 143; Indels 95; Gaps 18;  
QY 12 MRLLPSVLLGTAFTVYVLAAGVRLLSAFLPARFYQALDRLCYQSVLFEFFNYTG 71  
DB 20 MRFQFVYVNLVAIPSYICVYIIQLPLRLVDSKRFMY-IGIMYKMLGMYA-SMGVYAG 77  
QY 72 VOILLYGDLPK--NKEKITLANHSTVDWIVADILAIKRON--ALGHVRYVLKEGAKML 126  
DB 78 YTVMEWEDIKRANVSKDAVLAHQAATGD--VCTLMCLDQKGLVAMOMMLMDHITKYT 135  
QY 127 PLYGCEFAHQGIYVKSAAFKENE--MNRKLOSVDAGTPMYLVFPEGTRYNPQTK 183  
DB 136 N-EGIVSLVHGDFIRQGFYSRDQQLLLKKHLENNYRSRDKWIVLPEGGFLR---K 190  
QY 184 VLSAQAFQAQGLAVLKHVLPRIKTHVAFDCM---KN-----YLDL----- 224  
DB 191 RRETSQAFARKNNLPFLTNVTPRSGATKIILNLAQOKNSPAGDAKELSKSGKLG 250  
QY 225 -IYDVTVVYVGKDD-----GGORRESPMTTEFLCKECPRIHIHIDRIKDDPEDEH 276  
DB 251 WIDTTIATYAKAEPIIDQITILGTRK--PTVT-----AHYIIFPKDVPLETD 298  
QY 277 MRRMLHREFIKDKMLIEFYES--PDPERRRKRPKGSVNSKISIKTLPSMLISGITA 333  
DB 299 LTTWLVGREFVEKEDLSHFYETGAF-----PSKGHKEAVSREMTLS----- 340  
QY 334 GMLMTDAGRKLYVNTWT-----YGTLLGCLMWTI 362  
DB 341 -----NLWIFLIQSFAFLSGYMWYNI 361

RESULT 9  
YTT3\_CAEEL STANDARD; PRT; 391 AA.  
AC Q11087;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE Hypothetical 44.7 kDa protein COIC10.3 in chromosome X.  
GN COIC10.3.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
OC Rhabditidae; Peloderinae; (Caenorhabditis).  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISOL N2;  
RA Pauley A.;  
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: SOME, TO C.ELEGANS F0865.2.  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: U23526; AAB36850.1; -  
DR WormRep: COIC10.3; CE02449.  
DR InterPro: IPR002123; Acyltransferase.  
KW Hypothetical protein.  
SQ SEQUENCE 391 AA; 44718 MW; FB5C062A0172A8C1 CRC64;

Query Match 9.5%; Score 182; DB 1; Length 391;  
Best Local Similarity 25.0%; Pred. No. 4.9e-08;  
Matches 72; Conservative 47; Mismatches 105; Indels 64; Gaps 15;  
QY 69 YTGVOILLYGD--LPKNENIYLANHSTVDWIVADILAIKRONALGHVR---YVLKEG 122  
DB 98 FVGATVREYGTNLAGVAEERKCLLANHGLDHFV--LMQSLNGKSGISRMWVYINI 154  
QY 123 LKMLPLYGCFYAOHGCIYV-----KRSAKFN--EKERKLOSVDAGTPMYLVIFPEGT 175  
DB 155 WKYTPV-LGVMTSHGNFVGVGSKRDSVLSFPRDHLKNSFYKY-DYG---WVIMYPEG 209  
QY 176 RYVNEQTFKVLASQAFQAQGLAVLKHVLPRIKTHVAFD-----CMKNY- 221  
DB 210 RL-----YLVANSGRTPFEKNGKRLPDCVYPRGAAVAVLDVGLPTDLSMSKCGGEP 265  
QY 222 LDAIYDVTVVYEGKDDGGORRESP---TWTEFLCKECPRIHIHIDRIKDDPEDEH 276  
DB 266 IKYIIDATIGY-----RKGAVPDIDCVAMGQESVSAQFAVHYDVIPVKEWSDENL 318  
QY 277 MRRMLHREFIKDKMLIEFYESPDPERRRKRPKGSVNSKISIKTLPS 324  
DB 319 LKEFLERYIITKDLAEFYKT-----GHFPGDKT-----KVIPI 353

RESULT 10  
YIHG\_ECOLI STANDARD; PRT; 310 AA.  
ID YIHG\_ECOLI  
AC P32129;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein yihg.  
GN YIHG OR B3862.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K12 / MG1655;  
RX MEDLINE=93347969; PubMed=8346018;  
RA Plunkett G. III, Burland V.D., Daniels D.L., Blattner F.R.;  
RT "Analysis of the Escherichia coli genome. III. DNA sequence of the  
RT region from 87.2 to 89.2 minutes."  
RL Nucleic Acids Res. 21:3391-3398(1993).  
CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
CC ACYLTRANSFERASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: L19201; AAB02997.1; -

DR EMBL: AE000461; AAC76860.1; -.  
 DR PIR: S40808; S40808.  
 DR Ecogene: EG11833; y1ng.  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 310 AA; 36289 MW; 9F8E3F52EB0B186E CRC64;

Query Match 8.5%; Score 164.5; DB 1; Length 310;  
 Best Local Similarity 20.5%; Pred. No. 1.1e-06;  
 Matches 66; Conservative 69; Mismatches 128; Indels 59; Gaps 13;

QY 1 MLISVLHYHSMRYLLPSVVLGTAPTYVLAWGVWRLLSAFARYQALDRLCYVYQS 60  
 DB 19 LILSTVLTLTITFCGVPIITIGIVKLLPVPYIMKVSF-----CDFMYTCWCEG 70  
 QY 61 MWLFFENYTGVOILLYGDLPRKNEN-IYLANHSTVDWIADILAIKONALGHV---R 116  
 DB 71 LAVLHLN-PHLQWEVHGLEJLSKKNWYLLICNHRSMADIVLCVLFK-----H1PMK 124  
 QY 117 YVLKELKALPLYG--C-----YFAQHGITYKRSKAKNEKEMKRLKQSYVDAGT-PMY 167  
 DB 125 YELKQQLAVVPELGACWSLMDPEFMKRYRAYLLRHPERKGVETTRSCCKEFLHPPT 184  
 QY 168 LVIFPEGTYRNPFOFKVLSAQAFNAORGLAVLKHYLTPRIKATHVAFQCMKYLDAID 227  
 DB 185 IYNFVDSGKFTQEK-----HQOTHSTFQNLPPKAGIAMALNVLGKQPKLNL 233  
 QY 228 VTWVEYEGKDDGQRRSPMTTEFLCKECPKIHIDRIDKRVPE-----QEH 276  
 DB 234 VILCY--PDNNRQ-----PEFDMLSGKLTIRIVH---VDLPADLHGDIYDKSFKRH 283  
 QY 277 MRRWLHERFEIDKMLIEFYES 298  
 DB 284 FOQWLMSLWQEKDRLLTSLMS 305

## RESULT 11

PLSC\_YEAST STANDARD; PRT; 303 AA.  
 ID PLSC\_YEAST STANDARD; PRT; 303 AA.  
 AC P33333;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Probable 1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51)  
 DE (1-AGP acyltransferase) (LPAAT).  
 DE acyltransferase) (LPAAT).  
 GN SLC1 OR YDL052C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycos.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94012814; PubMed=8408076;  
 RA Nagiec M.M., Wells G.B., Lester R.L., Dickson R.C.;  
 RT "A suppressor gene that enables Saccharomyces cerevisiae to grow  
 without making sphingolipids encodes a protein that resembles an  
 Escherichia coli fatty acyltransferase."  
 RL J. Biol. Chem. 268:22156-22163(1993).  
 [2]  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA urrestarazu L.A., Andre B., Vissers S.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 182-303 FROM N.A.  
 RA Bloeker H., Brandt P.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: MAY BE AN ACYLTRANSFERASE WITH AN ALTERED SUBSTRATE  
 CC SPECIFICITY THAT ENABLES IT TO USE A C-26 COA IN PLACE OF THE  
 CC C-16 OR C-18 COAS USED BY THE WILD TYPE PROTEIN.  
 CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate =

CC COA + 1,2-diacyl-sn-glycerol 3-phosphate.  
 CC -1- PATHWAY: FIRST STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
 CC ACYLTRANSFERASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: L13282; AAA16514.1; -.  
 DR EMBL: Z74100; CAA98614.1; -.  
 DR PIR: A48600; A48600.  
 DR SGD: S0002210; SLC1.  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KW Phospholipid biosynthesis; Transferase; Acyltransferase;  
 KW Transmembrane.  
 FT TRANSMEM 16 30 POTENTIAL.  
 FT VARIANT 44 44 Q -> L (IN SUPPRESSOR ALLELE SLC1-1).  
 SQ SEQUENCE 303 AA; 33887 MW; 36CBBC2659655EB CRC64;

Query Match 6.4%; Score 124; DB 1; Length 303;  
 Best Local Similarity 27.6%; Pred. No. 0.003;  
 Matches 55; Conservative 28; Mismatches 78; Indels 38; Gaps 8;

QY 10 YSMRYLLPSVVLGTAPTYVLAWGVWRL-----LSAFPARFQALDRLCYVYQSMVLE 64  
 DB 9 YIRSVLVVIALAGCGFYGVYASITCTLGKQLHQAQWTR-----CTYHWKLM 58  
 QY 65 FFENYTGVOILLYGDLPRKNENIYLANHSTVDWIADILAIKONALGHV---RYVL 119  
 DB 59 L-----GLDYKVVGEENLAKKPYIMIANHQSTLDFM-----LGRIFPGCTVNA 103  
 QY 120 KEGLKALPLYGCTFAQHGITYKRSKAKNEKEMKRLKQSYVDAGTMYLVIFEGRR-TN 178  
 DB 104 KSLKLYVPELGMFALSGYFIDRSKROEALDTLNKLENVKR-NRRALWVPEGTRSYT 162  
 QY 179 PEQFKVLSAQAF-AAORG 196  
 DB 163 SELTMLPFKKGAFLAQOG 181

## RESULT 12

PLSC\_LIMDO STANDARD; PRT; 281 AA.  
 ID PLSC\_LIMDO STANDARD; PRT; 281 AA.  
 AC Q42870; Q40120;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE 1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51) (1-AGP  
 DE acyltransferase) (LPAAT).  
 DE (LPAAT).  
 GN PLSC.  
 OS Limnanthes douglasii (Douglas's meadowfoam).  
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Limnanthaceae; Limnanthes.  
 OX NCBI\_TaxID=28973;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96028122; PubMed=7588719;  
 RA Hanke C., Wolter F.P., Coleman J., Peterek G., Frenzen M.;  
 RT "A plant acyltransferase involved in triacylglycerol biosynthesis  
 RT complements an Escherichia coli sn-1-acylglycerol-3-phosphate  
 RL acyltransferase mutant."  
 RL Eur. J. Biochem. 232:806-810(1995).  
 [2]

```

RP SEQUENCE FROM N.A.
RX MEDLINE-96046746; PubMed-7579178;
RA Brown A.P., Brough C.L., Kroon J., Slabas A.R.;
RT Identification of a cDNA that encodes a 1-acyl-sn-glycerol-3-
RT phosphate acyltransferase from Limnantes douglasi.;
RL Plant Mol. Biol. 23:267-278(1995).
CC -1- FUNCTION: CONVERTS LYSPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC
CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION. THIS ENZYME
CC USES ERICOYL-COA AS AN ACYL DONOR.
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate =
CC CoA + 1,2-diacyl-sn-glycerol 3-phosphate.
CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE
CC ACYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X83266; CA58239.1; -
DR EMBL: 246836; CA86877.1; -
DR InterPro: IPR002123; Acyltransferase.
DR Pfam: PF01553; Acyltransferase: 1.
KM Phospholipid biosynthesis; Transferase; Acyltransferase;
KM Transmembrane.
FT TRANSMEM 40 60 POTENTIAL.
FT TRANSMEM 71 91 POTENTIAL.
FT TRANSMEM 110 130 POTENTIAL.
FT CONFLICT 46 46 I -> V (IN REF. 2).
FT CONFLICT 188 188 R -> G (IN REF. 2).
FT CONFLICT 262 262 V -> I (IN REF. 2).
FT CONFLICT 281 281 N -> K (IN REF. 2).
SQ SEQUENCE 281 AA; 31716 MW; 9C880BD9E492EE2A CRC64;

Query Match 6.2%; Score 119; DB 1; Length 281;
Best Local Similarity 23.8%; Pred. No. 0.0072;
Matches 65; Conservative 40; Mismatches 92; Indels 76; Gaps 14;

QY 13 RYLPSV-----VLGTAPTYV-----LAWGWRLLSAFLPARFOAL 50
QY 13 RYLPVAVATADDDKOSVFVLLSCFKIFCFALVLTAVAKG--IMVLLPFPYRIR 70
DB 13 RYLPVAVATADDDKOSVFVLLSCFKIFCFALVLTAVAKG--IMVLLPFPYRIR 70
QY 51 DRLRY-CVQSMVLFEPENTGVQIILYGLPKNKENIITLANHSTVDWIVADILAIQ 109
DB 71 LGLNLYGHITIGLVIMYV---GIPKIQGS-EHTKKRAIYISNHSPIDAFFVWMLA--- 122
QY 110 NALGHRYVLEKGLKMLPLGYCFAGHGGIYVRSKAFNKKENKQSVVDAGTP--MY 167
DB 123 -PIGTGVAKKEVY-WPLGLGULTTLANHTRIDRS--NPAALIQSKKEAVRITENKLS 177
QY 168 LVTFPESTRYNPQTKVLSAQAFA--AAORGAVLKLHVLTPIKATHVAF----- 215
DB 178 LIMEPESTR--SRDGRLLPCKKGFVHLALQSHLPYPMILT-----GTHLAMRGTFEVRP 231
QY 216 -----DCMKNYLDATYDTV 230
DB 232 VPITVRYLPINTDWTVDKIDYVKMIDHYV 264

RESULT 13
PLSC_MYCPN STANDARD; PRT; 266 AA.
AC P75479;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable 1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51)

```

```

DE (1-AGP acyltransferase) (1-AGPAT) (Lysophosphatidic acid
DE acyltransferase) (LPAT).
GN PLSC OR MPN299 OR MP537.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE-97105885; PubMed-8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -1- FUNCTION: CONVERTS LYSPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC
CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION.
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate =
CC CoA + 1,2-diacyl-sn-glycerol 3-phosphate.
CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.
CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE
CC ACYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000052; AAB96185.1; -
DR InterPro: IPR002123; Acyltransferase.
DR Pfam: PF01553; Acyltransferase: 1.
KM Phospholipid biosynthesis; Transferase; Acyltransferase;
KM Complete proteome.
SQ SEQUENCE 266 AA; 30414 MW; 79D933AD8203927F CRC64;

Query Match 6.0%; Score 116; DB 1; Length 266;
Best Local Similarity 25.2%; Pred. No. 0.012;
Matches 56; Conservative 43; Mismatches 75; Indels 48; Gaps 13;

QY 10 YSMRYL-LPSVYLTGTPYVILAMGWRLLSAFLPARFOALD-----RLCYVQSMV 62
QY 11 FCLRFQLDLSLVV--LPFVYL-----MLSSLSANYSISENYPPELRFKVVYLV 63
DB 11 FCLRFQLDLSLVV--LPFVYL-----MLSSLSANYSISENYPPELRFKVVYLV 63
QY 63 LFPFENTGVQIILYGLPKN--KENIITLANHSTVDWIVADILAIQNALGHVRYLK 120
DB 64 LFLY--IKGVKVVIVN--PENVRKAVLVVANHKSNDPLILKAFGTEGVPLPEIAK 119
QY 121 EGLK--WL-----PIYGYCFAGHGGIYVRSKAFNKKENKRL--OSVYAGTPMYLV 171
DB 120 IELODTWLFKIMKLIDVFID-----RNKLKRWMAASLEQOOIINOFTA--LCVF 167
QY 172 PEGTRYNPEQTKVLSAQAFAAGLAVLKLHVLTPIKATHV 213
DB 168 PEGTR-----VLSRQIGEFKSGALKVAVNAFVPIPLTV 202

RESULT 14
PLSC_LIMAL STANDARD; PRT; 281 AA.
AC Q42868;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51) (1-AGP
DE acyltransferase) (1-AGPAT) (Lysophosphatidic acid acyltransferase)
DE (LPAT).
OS Limnantes alba (White meadowfoam).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```



---



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:16:55 ; Search time 65.95 Seconds

(without alignments)  
954.817 Million cell updates/sec

Title: US-09-853-526-5  
Perfect score: 1924  
Sequence: 1 MLSTLVHTYSMRYLPSVY.....YVNTWYGTLLGLMTWTRKA 364

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_19:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1924	100.0	364	4 Q9BOG4	Q9BOG4 homo sapien
2	1868	97.1	353	4 Q9NUQ2	Q9NUQ2 homo sapien
3	1568.5	81.5	354	11 Q9D1E8	Q9D1E8 mus musculu
4	470	24.4	386	5 Q9SR12	Q9SR12 caenorhabdi
5	398.5	20.7	373	10 Q9SSH0	Q9SSH0 arabidopsis
6	398.5	20.7	373	10 Q9C9P8	Q9C9P8 arabidopsis
7	390.5	20.3	375	10 Q9LHNA	Q9LHNA arabidopsis
8	328	17.0	374	10 Q41745	Q41745 zea mays (m
9	309.5	16.1	390	10 Q9XFW4	Q9XFW4 brassica na
10	307	16.0	377	10 Q40119	Q40119 limnanthes
11	305.5	15.9	306	10 Q9SDN3	Q9SDN3 prunus dulc
12	272.5	14.2	386	5 Q9VVS1	Q9VVS1 drosophila
13	270	14.0	378	11 Q9Z4S1	Q9Z4S1 rattus norv
14	267	13.9	380	5 Q9VVA9	Q9VVA9 drosophila
15	263	13.7	376	10 Q9SYC8	Q9SYC8 arabidopsis
16	251	13.0	316	11 Q9DB84	Q9DB84 mus musculu

17	249.5	13.0	350	3 Q94361	Q94361 schizosacch
18	242.5	12.6	344	5 Q16526	Q16526 caenorhabdi
19	239	12.4	311	10 Q39317	Q39317 brassica na
20	238.5	12.4	310	10 Q9SVX9	Q9SVX9 arabidopsis
21	223	11.6	314	11 Q9D517	Q9D517 mus musculu
22	218.5	11.4	399	5 Q23087	Q23087 caenorhabdi
23	212	11.0	439	5 Q20800	Q20800 caenorhabdi
24	205.5	10.7	370	11 Q91YX5	Q91YX5 mus musculu
25	189.5	9.8	426	5 Q01882	Q01882 caenorhabdi
26	189.5	9.8	428	5 Q95R03	Q95R03 caenorhabdi
27	184	9.6	295	16 Q91657	Q91657 pseudomonas
28	153	8.0	304	16 Q915A6	Q915A6 pseudomonas
29	145	7.5	359	10 Q9FF57	Q9FF57 arabidopsis
30	131.5	6.8	363	5 Q19221	Q19221 caenorhabdi
31	130	6.8	294	16 Q9KVV5	Q9KVV5 vibrio chol
32	124	6.4	248	2 Q30849	Q30849 salmonella
33	116	6.0	281	10 Q9SD02	Q9SD02 limnanthes
34	114.5	6.0	247	16 Q9X219	Q9X219 thermotoga
35	113.5	5.9	385	5 Q9NFP1	Q9NFP1 leishmania
36	113	5.9	262	16 Q98G17	Q98G17 rhizobium 1
37	112	5.8	241	16 Q9CPE2	Q9CPE2 pasteurella
38	112	5.8	258	16 Q9HW50	Q9HW50 pseudomonas
39	109.5	5.7	235	2 Q9KX14	Q9KX14 rhodospirillum rubrum
40	109.5	5.7	488	12 Q9WT03	Q9WT03 human herpes
41	108	5.6	261	16 Q9KPE3	Q9KPE3 vibrio chol
42	105.5	5.5	249	2 Q9EY25	Q9EY25 xanthomonas
43	105	5.5	260	16 Q92MG4	Q92MG4 rhizobium m
44	104	5.4	344	10 Q9L1Y4	Q9L1Y4 brassica na
45	102.5	5.3	291	3 Q917C5	Q917C5 emeticella

## ALIGNMENTS

RESULT 1  
Q9BOG4 PRELIMINARY: PRT: 364 AA.  
AC Q9BOG4;  
DT 01-JUN-2001 (TRENBLREL. 17, Last sequence update)  
DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)  
DE HYPOTHETICAL 42.1 KDA PROTEIN (LYSOPHOSPHATIDIC ACID  
DE ACYLTRANSFERASE-EPSILON) (EC 2.3.1.51).  
GN DKFZP761C222 OR LPAAT-E.  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC TISSUE=AMGDALAE;  
RX MEDLINE=21154917; PubMed=11230166;  
RA Wiemann S., Well B., Wellenreuther R., Gassenhuber J., Glass S.,  
RA Ansgore W., Boecker M., Bloeker H., Baurerachs S., Blum H.,  
RA Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strick N.,  
RA Mewes H.W., Othenmaier B., Obermaier B., Tampe J., Heubner D.,  
RA Wambutt R., Korn B., Klein M., Poustka A.,  
RA "Towards a Catalog of Human Genes and Proteins: Sequencing and  
RT Analysis of 500 Novel Complete Protein Coding Human cDNAs";  
RN genome Res. 11:422-435(2001).  
RP (2)  
RP SEQUENCE FROM N.A.  
RA Leung D.W.;  
RT "Cloning and expression of LPAAT-epsilon";  
RT Submitted (May-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL136587; CAB66522.1; -;  
DR EMBL; AF375789; AAK54809.1; -;  
DR InterPro: IPR002123; Acyltransferase.  
DR Pfam: PF01553; Acyltransferase; 1.  
KW Hypothetical protein; Transferase; Acyltransferase.  
SQ SEQUENCE 364 AA: 42072 MW: 90A0F87FC7C8081 CRC64;

Query Match 100.0%; Score 1924; DB 4; Length 354;  
Best Local Similarity 100.0%; Pred. No. 1.8e-167;  
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLSTLVHTSMRYLLPSVLLGTAFTYVLAAGVWRLSAPLPAFYQALDRLCYCYOAVLFEFFENYTG 60  
DB 1 MLSTLVHTSMRYLLPSVLLGTAFTYVLAAGVWRLSAPLPAFYQALDRLCYCYOAVLFEFFENYTG 60  
QY 61 MVLFFENYTGVOILLTGDPKKNENIYLANHSTYDWMIVADILAIRONALGHVRYVVK 120  
DB 61 MVLFFENYTGVOILLTGDPKKNENIYLANHSTYDWMIVADILAIRONALGHVRYVVK 120  
QY 121 EGLKMLPLCYCTFAQHGIGYVKSAAKNEKEMRNKLOSVDAGTPMYLVIFPESTRNPE 180  
DB 121 EGLKMLPLCYCTFAQHGIGYVKSAAKNEKEMRNKLOSVDAGTPMYLVIFPESTRNPE 180  
QY 181 QTVLASQAFAAORGAVLKHVLTPIKATHVAFDCMKNYLDAIYDVTYVEGKDDGGQ 240  
DB 181 QTVLASQAFAAORGAVLKHVLTPIKATHVAFDCMKNYLDAIYDVTYVEGKDDGGQ 240  
QY 241 RRESPTEFLCKECPKIHIDRIDKDVPEEOEHMRMLHEFEIKDKMLIEFYESPD 300  
DB 241 RRESPTEFLCKECPKIHIDRIDKDVPEEOEHMRMLHEFEIKDKMLIEFYESPD 300  
QY 301 PERKKRPKGSVNSKLSIKKTPDSMLISGLTGAAGMLMTDAGRKYVNTWITGTLGLWY 360  
DB 301 PERKKRPKGSVNSKLSIKKTPDSMLISGLTGAAGMLMTDAGRKYVNTWITGTLGLWY 360  
QY 361 TIRKA 364  
DB 361 TIRKA 364

RESULT 2  
ID Q9NU02 PRELIMINARY; PRT: 353 AA.

AC Q9NU02; 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DE CDNA FLJ11210 FIS, CLONE PLACE1007954.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
NCBI\_Taxid=9606;  
RP SEQUENCE FROM N.A.  
RC TISSUE=PLACENTA;  
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
Wagatsuna M., Hosoliri T., Kaku Y., Kodaira R., Kondo H., Sugawara M.,  
Takatsuchi M., Chida Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y.,  
RA Minomiya K., Iwayanagi T.,  
RT "NEO human cDNA sequencing project."  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AK002072; BAA92069.1;  
DR InterPro: IPR002123; Acyltransferase.  
DR Pfam: PF01553; Acyltransferase; 1  
SQ SEQUENCE 353 AA; 40799 MW; 0A8A1A42361F1B14 CRC64;

Query Match 97.1%; Score 1868; DB 4; Length 353;  
Best Local Similarity 99.7%; Pred. No. 2.3e-162;  
Matches 352; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 12 MRYLLPSVLLGTAFTYVLAAGVWRLSAPLPAFYQALDRLCYCYOAVLFEFFENYTG 71  
DB 1 MRYLLPSVLLGTAFTYVLAAGVWRLSAPLPAFYQALDRLCYCYOAVLFEFFENYTG 60  
QY 72 VQILLTGDPKKNENIYLANHSTYDWMIVADILAIRONALGHVRYVVKGLKMLPLCYGC 131

DB 61 VQILLTGDPKKNENIYLANHSTYDWMIVADILAIRONALGHVRYVVKGLKMLPLCYGC 120  
QY 132 YFAOHGIGYVKSAAKNEKEMRNKLOSVDAGTPMYLVIFPESTRNPEOTKVASQA 191  
DB 121 YFAOHGIGYVKSAAKNEKEMRNKLOSVDAGTPMYLVIFPESTRNPEOTKVASQA 180  
QY 192 AAORGAVLKHVLTPIKATHVAFDCMKNYLDAIYDVTYVEGKDDGGORRESPTEFL 251  
DB 181 AAORGAVLKHVLTPIKATHVAFDCMKNYLDAIYDVTYVEGKDDGGORRESPTEFL 240  
QY 252 CKECSPKIHIDRIDKDVPEEOEHMRMLHEFEIKDKMLIEFYESPDPERKRPKGS 311  
DB 241 CKECSPKIHIDRIDKDVPEEOEHMRMLHEFEIKDKMLIEFYESPDPERKRPKGS 300  
QY 312 VNSKLSIKKTPDSMLISGLTGAAGMLMTDAGRKYVNTWITGTLGLWYIRKA 364  
DB 301 VNSKLSIKKTPDSMLISGLTGAAGMLMTDAGRKYVNTWITGTLGLWYIRKA 353

RESULT 3  
ID Q9DIE8 PRELIMINARY; PRT: 354 AA.

AC Q9DIE8; 01-JUN-2001 (Tremblrel. 17, Created)  
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
DE 1110013A05RIK PROTEIN.  
GN 1110013A05RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_Taxid=10090;  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=EMBRYO;  
RX MEDLINE=21085660; PubMed=11217651;  
RA Kawai J., Shinagawa A., Shidara K., Yoshino M., Itoh M., Ishii Y.,  
Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,  
Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
Schriml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli P., Barsh G.,  
Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamija M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitlaker C., Wilming L.,  
RA Wysshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsuhl S.,  
RA Hayashizaki Y.,  
RT "Functional annotation of a full-length mouse cDNA collection."  
RL Nature 409:685-690(2001).  
DR EMBL: AK003649; BAB22915.1;  
DR MGD: MGI:1915880; 1110013A05RIK.  
DR InterPro: IPR002123; Acyltransferase.  
DR Pfam: PF01553; Acyltransferase; 1  
SQ SEQUENCE 354 AA; 40943 MW; 447EE924B91E800 CRC64;

Query Match 81.5%; Score 1568.5; DB 11; Length 354;  
Best Local Similarity 81.4%; Pred. No. 5.3e-135;  
Matches 288; Conservative 39; Mismatches 26; Indels 1; Gaps 1;

QY 12 MRYLLPSVLLGTAFTYVLAAGVWRLSAPLPAFYQALDRLCYCYOAVLFEFFENYTG 71  
DB 1 MRYLLPSVLLGTAFTYVLAAGVWRLSAPLPAFYQALDRLCYCYOAVLFEFFENYTG 60  
QY 72 VQILLTGDPKKNENIYLANHSTYDWMIVADILAIRONALGHVRYVVKGLKMLPLCYGC 131

Db 61 VOILLYGDLPKKNENIYLANHSTVDWIVADMIAARQDALGHVRYVLKDKLWMLPLYGF 120  
QY 132 YFAOHGIYKRSKAFNEKEMRNKLOSVDAGTPMYLVIPEEGTRVPEQTKVLSAQAF 191  
Db 121 YFAOHGIYKRSKAFNEKEMRNKLOSVDAGTPMYLVIPEEGTRVPEQTKVLSAQAF 180  
QY 192 AARGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVVYEGKRDG-GORRSPMTTEF 250  
Db 181 AARGLAVLKHVLTLPRIKATHVAFDSKSLDAIYDVTVVYEGNEKSGKYNPSPMTTEF 240  
QY 251 LKCECKRIHIDRIDKDYPEDEHMRKWLHERFEIKDKMLIEFYESDPERRKRPFGK 310  
Db 241 LKQCKRKLHIFRIDIRNEPEDEHMKWLHERFEIKDRLLIEFYDSDPPERENKFPKG 300  
QY 311 SVNSKSIKKTLPMSLISGLTAGMLTDAGRLVNTWYTGTLGCLWTTIKA 364  
Db 301 SVNSRSLVKTLPSVLISGLTAGMLTMSGRKLYMGWTLYGLCLWTFVKA 354

RESULT 4  
Q95R12 PRELIMINARY; PRT; 386 AA.  
AC Q95R12; 01-DEC-2001 (Tremblrel. 19, Created)  
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE HYPOHETICAL 44.0 KDA PROTEIN.  
GN F28B3.9.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematozoa; Chromadorea; Rhabditida; Rhabditidae;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RX MEDLINE=99069613; PubMed=9851916;  
RA None;  
RT "Genome sequence of the nematode C. elegans: a platform for  
investigating biology. The C. elegans Sequencing Consortium.";  
RL Science 283:2012-2018(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RX Giesel C., Kramer J., Smith A.;  
RA "The sequence of C. elegans cosmid F28B3.";  
RL Submitted (May-1997) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RX Waterston R.;  
RA "Direct Submission.";  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF003136; AAK93853.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 386 AA; 43978 MW; A1E3020C201452A6 CRC64;

Query Match 24.4%; Score 470; DB 5; Length 386;  
Best Local Similarity 33.0%; Pred. No. 1,4e-34;  
Matches 131; Conservative 75; Mismatches 143; Indels 48; Gaps 14;

QY 1 MLSTLVHNYSMRYLTP---SVLLIGTAPTYVLAMGVNRLLSAFLPARFYQALDDRLYVC 56  
Db 1 MLSTLVHNYSMRYLTP---SVLLIGTAPTYVLAMGVNRLLSAFLPARFYQALDDRLYVC 56  
QY 57 VYOSMVLFFENTYGVQIILYGDLPK--NK---ENIYLANHSTVDWIVADILAIKRON 110  
Db 57 SVYRCLTFPENLSGVEIYILHGTNEEYVKTGRPENAVMISNQSVDWITIPMLAARHG 116  
QY 111 ALGH---VRYLKEGLKMLPLYGCFYAOHGIIYKRSKAFNEKEMRNKLOSVDAGTPMY 167  
Db 117 DQGNBOAFVRYMVKNSIHLVPMFGWYLFQHGIIYVRRFEGICAPVROLKMLNESDPYW 176

QY 168 LVIFPEGTRVPEQTKVLSAQAFPAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYD 227  
Db 177 LVIFPEGTRVPEQTKVLSAQAFPAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYD 235  
QY 228 VTVVY-EGKRDG-----QRRSPMTTEFIC--KECPRIHIDRIDK 267  
Db 236 VTVVYEGKRDG-----QRRSPMTTEFIC--KECPRIHIDRIDK 267  
QY 268 KQVPEDEHMRKWLHERFEIKDKMLIEFYESDPERRKRPFGKSVNSKSIKKTLPMSL 327  
Db 296 DEVPKALERTWTIERFTEKRIIDEFY-SEKPSGSLP-----CVPISQTLPSLTF 348  
QY 328 LS-GLTAGMLTDAGRLVNTWYTGTLGCLWTTIK 363  
Db 349 FSAALLAPFSRTIGR-IYLLTIASSPL-IAMLHTR 383

RESULT 5  
Q9SSH0 PRELIMINARY; PRT; 373 AA.  
AC Q9SSH0; 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE F25A4.2 PROTEIN.  
GN F25A4.2.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLOMBIA;  
RA Vysotskaia V.S., Schwartz J.R., Yu G., Toriumi M., Lenz C., Liu S.,  
Lee J.M., Li J., Gonzalez A., Liu A., Liu K., Vaysberg M., Sakano H.,  
Chin C., Choi E., Chou J., Altafi H., Araujo R., Brooks S.,  
Buehler E., Chao Q., Conn L., Conway A.B., Dunn P., Hansen N.,  
Hwang B., Huizar L., Khan S., Kim C., Palm C., Kowley D., Shinn P.,  
Walker M., Davis R.W., Ecker J.R., Federjpiel N.A., Theologis A.;  
RA "Arabidopsis thaliana chromosome 1 BAC F25A4 sequence.";  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AC008263; AAD55275.1; -.  
DR InterPro: IPR002123; Acyltransferase.  
DR Pfam: PF01553; Acyltransferase; 1.  
SQ SEQUENCE 373 AA; 42383 MW; 98642948D94709E6 CRC64;

Query Match 20.7%; Score 398.5; DB 10; Length 373;  
Best Local Similarity 29.8%; Pred. No. 4.7e-28;  
Matches 97; Conservative 68; Mismatches 117; Indels 43; Gaps 7;

QY 12 MRYLPSVLLGTAPTYVLAMGVNRLLSAFLPARFYQALDDRLYCYVQSMVT----- 63  
Db 23 LRGLMILLVPLSTAFMFLT-----YFAP---IAALGLRLISVQOSKRVSLIFGLM 70  
QY 64 -----FFPENYGVQIILYGDLPKKNENIYLANHSTVDWIVADILAIKRONLGHVRY 118  
Db 71 LALMPLTFEYVNTVTVFSGDILPVEKRVLLIANHREVDWMLMIALRKGLGIYIKYV 130  
QY 119 LRGLMILLVPLSTAFMFLT-----YFAP---IAALGLRLISVQOSKRVSLIFGLM 178  
Db 131 LRGLMILLVPLSTAFMFLT-----YFAP---IAALGLRLISVQOSKRVSLIFGLM 190  
QY 179 PEQTKVLSAQAFPAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVVYEGKRDG 238  
Db 191 EKKCK---RSQKRAAEVGLPALSNVLLPKTRGVGVCEVYLNHSLDAVYDITAYKPR--- 244  
QY 239 GORRESPTMTTEFICKECP-KIHITHRIDKDYPEDEHMRKWLHERFEIKDKMLIEFYE 297  
Db 245 -----CPSEFDNVEGTDPSEVNIHVRVYLLKEIPANEAESSAWLMDSFRIKDLISDF-- 297  
QY 298 SPDEPKRRKRPFGKSVNSKSIKKTLL 322

Db 298 ----NAQKFPNORPEEELSVLKCI 318

RESULT 6  
ID Q9C9P8 PRELIMINARY; PRT; 393 AA.  
AC Q9C9P8; 01-JUN-2001 (Tremblre, 17, Created)  
DT 01-JUN-2001 (Tremblre, 17, Last sequence update)  
DE PUTATIVE ACYL-COA:1-ACYLGlycerol-3-PHOSPHATE ACYLTRANSFERASE, 31588-29381.  
GN F9E10.13.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. COLUMBIA;  
RA MEDLINE=21016719; PubMed=11130712;  
RA Theologis A., Ecker J.F., Palm C.J., Federspiel N.A., Kaul S., White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y., Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W., Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Tewar K., Dunn P., Elgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y., Gall J.E., Goldsmith A.J., Haas B., Hansen N.F., Hughes B., Hultzer L., Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E., Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kvan A., Lam B., Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P., Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marzalli A., Mollis-Scher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I., Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D., Sakano H., Salerg S.L., Schwartz J.R., Shinn P., Southwick A.M., Sun H., Tallon L.J., Tambunga G., Tortum M.J., Town C.D., Utechtack T., Van Aken S., Vaysberg M., Vysotskaya V.S., Walker M., Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;  
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis thaliana."  
RL Nature 408:816-820(2000).  
EMBL: AC013258; AAG51931.1; -  
DR InterPro: IPR002123; Acyltransferase.  
DR Pfam: PF01553; Acyltransferase; 1.  
KW Transferase; Acyltransferase.  
SQ SEQUENCE 393 AA; 44892 MW; FA5949D21E8E9A29 CRC64;

Query Match 20.7%; Score 398.5; DB 10; Length 393;  
Best Local Similarity 29.8%; Pred. No. 5.1e-28;  
Matches 97; Conservative 68; Mismatches 117; Indels 43; Gaps 7;

Db 12 MRYLPSVLLGTAPRYVLANGVWRLSAFLPAFYQALDRLCYVQSNVL----- 63  
Db 23 LRGIMILVLSLSTAFMFL-----YFAP---IALGLRLSVQSRKVVSLIFGLW 70  
Qy 64 -----FFENVNGVQILLYGCLPRKKNIIYLANHOSVDMIVADIAIRONALGHRYV 118  
Db 71 LALMPYLFETVNGTIVVSGDIIPVEKRVLLIANHRTEDVMYIMNALRKGCGIYKYV 130  
Qy 119 LKESLKLPLKCYFAOHGIIYKRSAKFNEKEMRNKLOSYVDAGTPMYLVIFPEGRYRN 178  
Db 131 LKSLMLPLTFGWFHLEFIPVERKREVDPRVLQMLSSFKDQDEPLMLALFPEGTDFI 190  
Qy 179 PEQIKVLSAQAFAAAGGLAVLKHVLPRIKATVAFDCAKNYDAIDYDVTVYVEGDG 238  
Db 191 EENCK---RSQKFAAEVGLPALSNVLLPKTRGFGVCLVILANSIDAVYDLTIAKPR-- 244  
Qy 239 GQRSEPTMTFLCKECP-KIHIIHDIRDKKDVPEQEHHRMRLHEFEIKDKMLIFEYE 297  
Db 245 -----CSPFMDNVFGTDPSEVHIHVRVLLKEIPANEAESSAMLMDSFKLKDLSDF-- 297

Qy 298 SPDEPRKRRPFGKSVNSKLSIKRTL 322  
Db 298 ----NAQKFPNORPEEELSVLKCI 318

RESULT 7  
ID Q9LHN4 PRELIMINARY; PRT; 375 AA.  
AC Q9LHN4; 01-OCT-2000 (Tremblre, 15, Created)  
DT 01-OCT-2000 (Tremblre, 15, Last sequence update)  
DE 01-DEC-2001 (Tremblre, 19, Last annotation update)  
DE ACYL-COA:1-ACYLGlycerol-3-PHOSPHATE ACYLTRANSFERASE.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-COLUMBIA;  
RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;  
RL Submitted (May-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-COLUMBIA;  
RX MEDLINE=20363099; PubMed=10907853;  
RA Nakamura Y.;  
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II. Transferase features of the regions of 4,251,695 bp covered by ninety p1, TNG and BAC clones."  
RL DNA Res. 7:217-221(2000).  
EMBL: AP002039; BAB03094.1; -  
DR InterPro: IPR002123; Acyltransferase.  
DR Pfam: PF01553; Acyltransferase; 1.  
KW Transferase; Acyltransferase.  
SQ SEQUENCE 375 AA; 43490 MW; 8F6F5DC313A3D704 CRC64;

Query Match 20.3%; Score 390.5; DB 10; Length 375;  
Best Local Similarity 26.9%; Pred. No. 2.6e-27;  
Matches 101; Conservative 78; Mismatches 128; Indels 69; Gaps 9;

Qy 12 MRYLPSVLLGTAPRYVLANG-----VWRLSAFLPAFYQALDRLCYVQSNVLFFF 66  
Db 18 LRGICIMLVLSLSTAFMFLIFWGLSAVVLRLFS---IRYSRCVSPFFGSMALMPFLF 73  
Qy 67 ENTNGVQILLYGCLPRKKNIIYLANHOSVDMIVADIAIRONALGHRYVLEKGLKWL 126  
Db 74 EKINKTVIFSGDVPCEDVNLIANHRTEDVMYIMNALRKGQIGINIKYVLKSSIMKL 133  
Qy 127 PLKCYFAOHGIIYKRSAKFNEKEMRNKLOSYVDAGTPMYLVIFPEGRYRNPEQTKVLS 186  
Db 134 PLGMAHNLHEFIPVERKREVDENALQIYSPKDRPDALMLALFPEGTYTEAKCO--- 190  
Qy 187 ASQAFAAAGGLAVLKHVLPRIKATVAFDCAKNYDAIDYDVTVYVEGKDGQGRSEPT 246  
Db 191 RSKKFAENCLPILNVLPRTRGFGVSCLOELSCSDAVYDVTVIGYKTR----- 239  
Qy 247 MTERLCKECP-----KIHIIHDIRDKKDVPEQEHHRMRLHEFEIKDKMLIFE 234  
Db 240 -----CSPFMDNVYGIPESEVHIHRIINILQPNQKIDINAMLMNTFOLKDLND 291  
Qy 295 FYESPDEPRKRRPFGKSVNSKLSIKRTLPLMLISGLT-----AGMLMTDAGRKLX 345  
Db 292 FYSN-----GHRPNEGTEKEFPNTKYLINCLAVIAFTTICHTLPFPSSIMWF-----RIT 341  
Qy 346 VNTWIVYGLLGLMVT 361  
Db 342 VS-----LACVYLT 350

RESULT 8

Q41745  
ID Q41745 PRELIMINARY; PRT: 374 AA.  
AC Q41745;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)  
DT 01-NOV-1999 (TREMBlrel. 19, last annotation update)  
DE 1-ACYL-GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE (PUTATIVE).  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
OC Panicoideae; Andropogoneae; Zea.  
OX NCBI\_TaxID=4577;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BLACK MEXICAN SWEET; TISSUE-ENDOSPERM;  
RX MEDLINE=95035993; PubMed=7948871;  
RA Brown A.P., Coleman J., Tomney A.M., Watson M.D., Slabas A.R.;  
RT "Isolation and characterization of a maize cDNA that complements a 1-  
acyl-sn-glycerol-3-phosphate acyltransferase mutant of E.coli and  
RT encodes a protein which has similarities to other acyltransferases.";  
RL Plant Mol. Biol. 26:211-223(1994).  
DR EMBL, Z29518; CAA82638.1; -.  
DR InterPro: IPR002123; Acyltransferase.  
DR InterPro: IPR003015; HLH\_Myc.  
DR Pfam: PF01553; Acyltransferase; 1.  
DR PROSITE: PS00038; HELIX\_LOOP\_HELIX; UNKNOWN\_1.  
KW Transferase; Acyltransferase.  
SQ SEQUENCE 374 AA; 42571 MW; F1F5492CAFF24F93 CRC64;

Query Match 17.0%; Score 328; DB 10; Length 374;  
Best Local Similarity 29.8%; Pred. No. 1,3e-21;  
Matches 92; Conservative 64; Mismatches 121; Indels 32; Gaps 8;

QY 29 VLAMGVWLISA-----FLPAR-----FYQALDRLYCYQSWLFEFFENTGYQ 73  
DB 9 VLPLLLFLSLGLVNAIDQAVLFTVIRPFSKSFYRINRFLMLQLVWVDMWAGYK 68  
QY 74 ILLYD-----LPKNKENTITLANHOSTVDVITADILATRONALGHVRYLKEGLKWLPLX 129  
DB 69 VOLHADEETYSRSMGKEHALIISNHRSDIDWLIGWILAQSGCLGSLTAVMKSSKFLPYI 128  
QY 130 G--CYFAOHGIGYVRSKAFENKERNKLOSIVDAGTPMYLVIPEGTNRPQRTVLSA 187  
DB 129 GWSMFAEY--LFLERSNAKDEKTLKMGQLKDEPRFWLALFVEGRTFTPAK---LLA 183  
QY 188 SOAFAAQRGLAVLKHVLPRIKATVAFDCMKNYLDATVTVYVEGKDGQRRSPTM 247  
DB 184 AQEYAAASOGLPAPRVNLLPRTKGFSVSAVSIMRDFVPAIDTVIIVP-KDS-----PQPTM 237  
QY 248 TFFLKECPKTHIHIDRIKDVPEQEHMRRLHERFEIKMKMLIEFESPDPERKRF 307  
DB 238 LRIILGQSSVITHVRMKRHAEMSEPKSDEBVSCKDIFVAKALLDKHLATGTFTDEIRP 297  
QY 308 PEGSVNSKL 316  
DB 298 IGRPVKSL 306  
RESULT 9  
Q9XFW4  
ID Q9XFW4 PRELIMINARY; PRT: 390 AA.  
AC Q9XFW4;  
DT 01-NOV-1999 (TREMBlrel. 12, Created)  
DT 01-NOV-1999 (TREMBlrel. 12, last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)  
DE ACYL-COA:1-ACYLGLYCEROL-3-PHOSPHATE ACYLTRANSFERASE (EC 2.3.1.51).  
GN LPAAT.  
OS Brassica napus (Rape).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Brassica.

OX NCBI\_TaxID=3708;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-DEVELOPING SILIQUES;  
RA Graefin zu Muenster A., Wolter F.P., Frentzen M.;  
RT "A cDNA encoding a microsomal 1-acylglycerol-3-phosphate  
RT acyltransferase of Brassica napus L."  
RL Submitted (MAY-1997) to the EMBL/Genbank/DBJ databases.  
DR EMBL, Z295637; CAB09138.1; -.  
DR InterPro: IPR002123; Acyltransferase.  
DR InterPro: IPR003015; HLH\_Myc.  
DR Pfam: PF01553; Acyltransferase; 1.  
DR PROSITE: PS00038; HELIX\_LOOP\_HELIX; UNKNOWN\_1.  
KW Transferase; Acyltransferase.  
SQ SEQUENCE 390 AA; 43771 MW; F1446ELB3009C37 CRC64;

Query Match 16.1%; Score 309.5; DB 10; Length 390;  
Best Local Similarity 26.5%; Pred. No. 6,8e-20;  
Matches 83; Conservative 75; Mismatches 118; Indels 37; Gaps 9;

QY 69 YRGVQILLYGD-----LPKNKENTITLANHOSTVDVITADILATRONALGHVRYLKEGLK 124  
DB 63 MAGVKIQVFADETFENRMGKEHALVVCNHRSDIDWLIGWILAQSGCLGSALAVMKSSK 122  
QY 125 WPLVIG--CYFAOHGIGYVRSKAFENKERNKLOSIVDAGTPMYLVIPEGTNRPQRT 182  
DB 123 FLPLVIGSMWFSEY--LFLERNMAKDESTLQSGLQRLNDFPRFWLALFVEGRTFTPAK- 179  
QY 183 KYLSASQAFAAQRGLAVLKHVLPRIKATVAFDCMKNYLDATVTVYVEGKDGQRR 242  
DB 180 --LKAQAEYAAASELPVPRNVNLLPRTKGFSVSAVSNRSPVPAIDTVIIVP-----KTS 231  
QY 243 EEPYTFELCKCPKTHIHIDRIKDVPEQEHMRRLHERFEIKMKMLIEFESPDPE 302  
DB 232 PPPTMLRLFGQPSVYVHVHAIKCHSMKDLPEPDEIAIQMCRDQFVAKDALDKHIAA---- 287  
QY 303 RRRPRGKSVNSKLSIKTKLPLMLIISG---LTAGML-----MTDGRKLYVNTWY 351  
DB 288 --DTPFGQ--KEQNTGRPKSLAVVSWACLLITGAMKFLHWSNLFSSKKGIASAFGL 342  
QY 352 GTLLGCLWYTIKA 364  
DB 343 GIITLCMQILINS 355

RESULT 10  
Q40119  
ID Q40119 PRELIMINARY; PRT: 377 AA.  
AC Q40119;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)  
DE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE (PUTATIVE).  
OS Limnathes douglasii (Douglas's meadowfoam).  
OC Limnathes douglasii (Douglas's meadowfoam).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Limnathaceae; Limnathes.  
OX NCBI\_TaxID=28973;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC MEDLINE=96046746; PubMed=7579178;  
RX Brown A.P., Brough C.L., Kroon J., Slabas A.R.;  
RT "Identification of a cDNA that encodes a 1-acyl-sn-glycerol-3-  
RT phosphate acyltransferase from Limnathes douglasii.";  
RL Plant Mol. Biol. 29:267-278(1995).  
DR EMBL, Z48730; CAA88620.1; -.  
DR InterPro: IPR002123; Acyltransferase.  
DR InterPro: IPR003015; HLH\_Myc.  
DR Pfam: PF01553; Acyltransferase; 1.  
DR PROSITE: PS00038; HELIX\_LOOP\_HELIX; UNKNOWN\_1.  
KW Transferase; Acyltransferase.  
SQ SEQUENCE 377 AA; 42780 MW; 0DECDEF25D39687C CRC64;

Query Match 16.0%; Score 307; DB 10; Length 377;  
Best Local Similarity 25.5%; Pred. No. 1,1e-19;  
Matches 91; Conservative 85; Mismatches 131; Indels 50; Gaps 13;

QY 37 LLSAFLPARY---QALDDRLCYCYOSM-----LFFEPENTGQIILYGLPR--- 82  
DB 21 LVNFIQAVEFYVLRPISKDTYRINTLVALLMLLVWIDWAGKAVOLDYDTESFRL 80  
QY 83 -NKENITYLNHSTQVMIADILAIRONALGHVRYLKEGLKMLPLYG--CYFAOHGI 139  
DB 81 MGEHALLICNHRSDIDMLGVLAQRGCLSSIAVMKSSKFLPYIGMSMIFSE--L 138  
QY 140 YKRSKAFNEKEMRNKLSQSYVDAGTPLYVIFPESTRYNEPQTKV-LSASQAFPAQRGLA 198  
DB 139 FLENNMAKDEKTLKSCIQRLNDFPKFPMALFVEGTRF---TKAKLLAAGEVNASGLP 194  
QY 199 VLKHVLTFRKATIVA-DCCKKNYLDATYDVYVEGKDGQRRSESTPTMEFLCKECPKI 258  
DB 195 VPRNVLLPRTKGFVSANMRSFPAIYDLTVALP-----KTTEQPTMLRLEFRKSSVY 248  
QY 259 HIHDIRDKDQVPEOIHMRHREFEIKDKMLIEFYESPDPERRKRPFGKSVNSKLSI 318  
DB 249 HHLKRLMKDLPKTDICVAKMKDQFISKD-ALLDKHVAED-----TFSGLEVYQ---DI 299  
QY 319 KTLPLSMILSG-----LTAGML-----MTDAGRKLVTWTWYGLLGLCLWYTKA 364  
DB 300 GRPKMSLVVVVSMWMLCLGLGVKRLQMSALLSSMKGMITTFVGLGYTVLMLILRS 356

RESULT 11  
Q9SDN3 PRELIMINARY; PRT; 306 AA.  
AC Q9SDN3; (Tremblrel. 13, Last sequence update)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE.  
OS Prunus dulcis (Almond) (Prunus amygdalus).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.  
OX NCBI\_TaxID=3735;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Campalans A., Pages M., Messeguet R.;  
RT "Identification of differentially expressed genes during dehydration  
in almond (Prunus amygdalus) using the cDNA-AFLP technique.";  
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF213937; AAF20003.1; -  
DR InterPro: IPR002123; Acyltransferase.  
DR InterPro: IPR003015; HLH\_Myc.  
DR Pfam: PF01553; Acyltransferase; 1.  
DR PROSITE: PS00038; HELIX\_LOOP\_HELIX; UNKNOWN\_1.  
KW Transferase; Acyltransferase.  
SQ SEQUENCE 306 AA; 33993 MW; CBD334496E25908D CRC64;

Query Match 15.9%; Score 305.5; DB 10; Length 306;  
Best Local Similarity 31.9%; Pred. No. 1,1e-19;  
Matches 80; Conservative 55; Mismatches 101; Indels 15; Gaps 7;

QY 84 KENITYLANHSTQVMIADILAIRONALGHVRYLKEGLKMLPLYG--CYFAOHGIYV 141  
DB 3 KEHALVLSNHRSDIDMLGVLAQRGCLSSIAVMKSSKFLPYIGMSMIFSE--LFL 60  
QY 142 KRSKAFNEKEMRNKLSQSYVDAGTPLYVIFPESTRYNEPQTKVLSASQAFPAQRGLAVLK 201  
DB 61 ERSWAKDEGLTKSGVQRLEKPOFPMALFVEGTRF--QAKLL-AAQEVAAATGALPVP 117  
QY 202 HVLTPRKATHVAFDCCKNTLDAIYDVYVEGKDGQRRSESTPTMEFLCKECPKIH 261

DB 118 NVLIPRTKGVTVASQKRSAPAIYDTVAIP-----KSSPAPMLRLEGRPSVYHV 171  
QY 262 IDRIDKADVEEEOHMRMLHREFEIKDKMLIEF-YESPDPERRKRPFGKSVNSKLSIKK 320  
DB 172 IKRHVMDLPETDQBAVAKMDIFVAKDALLDKHTVQGTGDDQLKVTGAPLKS-LLVYT 230  
QY 321 TLPSMLISGL 331  
DB 231 AMACLLTIGAL 241

RESULT 12  
Q9VVS1 PRELIMINARY; PRT; 386 AA.  
AC Q9VVS1; Q9VVS0;  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE CC4729 PROTEIN.  
GN CC4729.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).  
RC STRAIN-BERKELEY;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
RA Wen K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Aghayani A., An H.-J., Andrews-Frankkoch C., Baldwin D.,  
RA Balieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Betman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Borchan M.R., Bouck J., Brokstein P., Brotlier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glogek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegami C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Slier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Sytkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A  
CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.  
DR EMBL; AE003527; AAF49471.1; -  
DR EMBL; AE003527; AAF49472.1; -  
DR FlyBase; FBgn0036623; CC4729.  
DR InterPro: IPR002123; Acyltransferase.



```
QY 36 RLTSAPFAPFYQALDRKYCYQSMVLFPEFENTGVQIILYGDLPKN-----KENIIT 89
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 43 RSLAYVLHVSFYCI---LVCV-----AEMVAGSKLRYIID-PODECKFTGKEHGL 89
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 90 LANHOSVVDVIMVADILAI3QNALGHVRYVLKEGKMLPLKGCYFAOAGGIVYKRSAPENE 149
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 90 LMHTYETIDWLTMMITDKNLGKTKAYAKKMLRYVPLGWWMMMAEFPLDNPEKDK 149
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 150 KEMRNKLOSVDAGTPMYIYIFPEGTRYNPEQTKVLSASQAFQAORGLAVLKHYLPRIK 209
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 150 VVIKTQKEVESPDPYMLLLNAGTRFPAKHEL---SVKPAERGLPLKHLHLIRTK 206
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 210 ATHVAFDCMKNYDAIVYVYEGKDDGQRRBSPMTPELKECPKIHIDRIKKD 269
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 207 GFTTSLPTMKGICPAIDIMLAFKNAE-----PKPTLSQLNGEPVPEYIRYPLDV 261
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 270 VPPEOEHMRWMLHERFEIKOKMLIEFESDPERR---KRFGKSVNSKLSIKKTLPSML 326
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 262 VPDEKFAAAMMODFEFAEKDKITDSFHETGSFRNSGVKEVPEKIYKPRLS---TLNPL 318
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 327 -----ILSGLTNG 334
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 319 GMATFAVLCTLHYLTSLVNG 339
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
RESULT 15
Q9SYC8      PRELIMINARY;      PRT;      376 AA.
AC Q9SYC8:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PUTATIVE ACYL-COA:1-ACYLG-CYCEROL-3-PHOSPHATE ACYLTRANSFERASE.
GN F1M15.12.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid 1; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Federspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
RA Altafli H., Araujo R., Huizar L., Rowley D., Buehler E., Dunn P.,
RA Gonzalez A., Kremenevskala I., Kim C., Lenz C., Li J., Liu S.,
RA Lueros S., Schwartz J., Shinn P., Toriumi M., Vysotskaya V.S.,
RA Walker M., Yu G., Ecker J., Theologis A., Davis R.W.,
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC006085; AAD30638.1; -.
DR InterPro; IPR002123; Acyltransferase.
DR InterPro; IPR003015; HLH_Myc.
DR Pfam; PF01553; Acyltransferase; 1.
DR PROSITE; PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
KW Transferase; Acyltransferase.
SQ SEQUENCE 376 AA; 43435 MW; 0B9AE143B09EDAD0 CRC64;
```

Query Match 13.7%; Score 263; DB 10; Length 376;  
Best Local Similarity 23.0%; Pred. No. 1..1e-15;  
Matches 72; Conservative 76; Mismatches 117; Indels 48; Gaps 8;

```
QY 62 VLEFFENYTVQVILYGDLPK---NKENIITLANHOSVVDVIMVADILAI3QNALGHVRY 117
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 57 LILVLFDMWMAICIKINLYVDATLELIGKEHALVLSNHRSDIDWLTGWMMAQROVGLGSSLA 116
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 118 VLKEGLKMLPLYG--CYFAOHGCIYVYKRSAPFNEKERKNKLOSVDAGTPMYIYIFPEGT 175
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 117 IMKREAKYLPITIGSMWFSY--IFLERSNAKDENTILKAGFKLEDPMTFWLALFVEGT 174
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 176 RYNPEQTKVLSASQAFQAORGLAVLKHYLPRIKATHVAFDCMKNYDAIVYVYEGK 235
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 175 RFTQEK---LEAAOEYASIRSLPSRNVLPRIKGFSAVSEIRSFVPALYDCLTVH-- 229
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 236 DDGGQRRBSPMTPELKECPKIHIDRIKKDVPEOEHMRWMLHERFEIKDKMLIEF 295
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
DB 230 -----NNOPTPTLLRMFSGSSSEINLQWRHRKMSLPEPTDGIQWCODLEFTRDAOLEKY 285
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 296 YESDPERRRRRPFKGSVNSKLSI-----KTLPSMLISGLTACMLMTDAGRKLXYVTWIIY 351
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 286 FT-----KQVFSDEVHQINRPIKPLIYI-----IMLGELVF 318
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 352 GTLLGCLMTYIKA 364
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 319 GFELQLQWLISVA 331
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: August 28, 2002, 11:16:56  
Job time: 560 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:18:44 ; Search time 75.33 Seconds

(without alignments)  
336.185 Million cell updates/sec

Title: US-09-853-526-70

Perfect score: 228

Sequence: 1 MRYLLPSVLLGTAPRYVLA.....NYLDAYDVTVVEKGKDDG 228

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched:

Word size : 0

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A.Geneseq\_032802:\*

- 1: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:\*
- 2: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:\*
- 3: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:\*
- 4: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:\*
- 5: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:\*
- 6: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:\*
- 7: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:\*
- 8: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:\*
- 9: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:\*
- 10: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT:\*
- 11: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:\*
- 12: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:\*
- 13: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:\*
- 14: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:\*
- 15: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:\*
- 16: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:\*
- 17: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:\*
- 18: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:\*
- 19: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:\*
- 20: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:\*
- 21: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:\*
- 22: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	228	100.0	228	20	AAV36740	Truncated PGI prot
2	228	100.0	352	20	AAV36728	Human PGI protein
3	228	100.0	353	22	AAAB39591	Human polypeptide
4	228	100.0	353	22	AAAB85532	Human secreted pro
5	228	100.0	353	22	AAAB85552	Human secreted pro
6	228	100.0	364	20	AAV36729	Human PGI protein
7	228	100.0	364	22	AAAM50128	Human acyltransfer
8	228	100.0	364	22	AAU00665	Human lysophosphat
9	185	81.1	185	20	AAV36753	PGI splice variant
10	185	81.1	315	20	AAV36751	PGI splice variant
11	184	80.7	300	20	AAV36752	PGI splice variant

12	154	67.5	182	20	AAV36750	PGI splice variant
13	144	63.2	269	22	AAE09653	Human gene 15 enco
14	144	63.2	353	22	AAAB93595	Human protein sequ
15	127	55.7	372	22	AAAM41377	Human polypeptide
16	105	46.1	291	20	AAV36744	PGI splice variant
17	85	37.3	96	20	AAV36749	PGI splice variant
18	74	32.5	261	20	AAV36745	PGI splice variant
19	70	30.7	257	22	ABG01360	Novel human diagno
20	70	30.7	980	22	ABG11977	Novel human diagno
21	62	27.2	66	20	AAV36748	PGI splice variant
22	62	27.2	68	20	AAV36747	PGI splice variant
23	62	27.2	77	20	AAV36742	PGI splice variant
24	62	27.2	90	20	AAV36746	PGI splice variant
25	44	19.3	238	20	AAV36743	PGI splice variant
26	44	19.3	97	22	ABG12868	Novel human diagno
27	44	19.3	1104	22	ABG12107	Novel human diagno
28	32	14.0	354	20	AAV36741	Mouse PGI protein
29	9	3.9	9	20	AAV36734	PGI box 3 motif.
30	9	3.9	157	22	AAAM50130	Lysophosphatidic a
31	9	3.9	237	18	AAAM55540	H. pylori ORF 12ge
32	9	3.9	240	19	AAAM55540	H. pylori GHP 903
33	9	3.9	242	18	AAAM55441	H. pylori ORF 06gp
34	8	3.5	9	20	AAV36737	Acyl glycerol tran
35	8	3.5	239	22	ABBS2710	Escherichia coli p
36	8	3.5	396	21	AAV99484	Yeast acyltransfer
37	8	3.5	1820	22	ABBS2370	Drosophila melanog
38	7	3.1	12	16	AAAR87705	Cocount AG3PAT try
39	7	3.1	12	20	AAV43149	Cocount 1-acylglyc
40	7	3.1	65	16	AAAR87720	Cocount LPAAT clon
41	7	3.1	65	20	AAV43164	1-acylglycerol -3-p
42	7	3.1	82	16	AAAR87719	Cocount LPAAT clon
43	7	3.1	82	16	AAAR87719	Cocount LPAAT clon
44	7	3.1	82	20	AAV43160	1-acylglycerol -3-p
45	7	3.1	82	20	AAV43163	1-acylglycerol -3-p

#### ALIGNMENTS

RESULT 1

AAV36740 standard; Protein; 228 AA.

AAV36740;

27-SEP-1999 (first entry)

Truncated PGI protein sequence.

PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;

KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.

XX

OS Homo sapiens.

XX

PN W0932644-A2.

XX

PD 01-JUL-1999.

XX

PF 22-DEC-1998; 98WO-1B02133.

XX

PR 09-SEP-1998; 98US-009658.

XX

PR 22-DEC-1997; 97US-0996306.

XX

PA (GEST ) GENSET.

PI Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;

XX

DR N-PSDB; AA00926.

XX

PT Use of a prostate cancer associated gene and biallelic markers

XX derived from it

PS Claim 7; Page 208; 385pp; English.

The invention relates to a mammalian Pci gene and protein, and a set of Pci biallelic markers. The Pci polynucleotide and biallelic markers are used in a hybridisation assay, a sequencing assay, or in an allele-specific amplification assay for determining the identity of a nucleotide at a Pci-related biallelic marker. The methods can be used to detect and to assess the risk of developing cancer or prostate cancer. Early-stage diagnosis of prostate cancer relies on prostate specific antigen (PSA) dosage. However, the effectiveness of this is limited due to its inability to discriminate between malignant and non-malignant affections of the organ. A need exists for both a reliable diagnostic procedure which would enable early-stage diagnosis, and for preventative and curative treatments of the disease. The Pci gene can be used for detection of prostate cancer, and the risk of developing it in the future, and can also be used to determine therapies for the disease.

**SQ Sequence 228 AA;**

Query Match	100.0%;	Score 228;	DB 20;	Length 228;
Best Local Similarity	100.0%;	Pred. No. 2.1e-222;		
Matches 228;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	1	MRLLPSVVLGGATATYVLANGVWRLLSAPLPARTYQALDRLKVCVOSMYLFFENNTG	60
Db	1	mrllpsvvlvggattatylangvwrllsalfpartfygaldrlkvcvysmwlffenytg	60
Qy	61	VOILLGDLPRKKKEIITLANQNSVDMIVADILAIRONALGHVRYVLKEGKATWLZLGC	12
Db	61	vqlllygdlprkknelltlanngsvdmivadilairnalgnyvylkegikwlpilzgc	12
Qy	121	YFAOHGIGTYKRSANAFNKENKRNKLQASVVDAGTPMYLIFPEGSTRYNEPQTFVLSASQAF	18
Db	121	yfaohgigtvkrsafnkekennkrlqgsyvdsagtpmylilfpegtrtynpeqctvlssaqaf	18
Qy	181	AAQGLAVLKHVLPRIKATHVADCKMKNYLDATYDVTVYEGKDDG	228
Db	181	aagglavlkhnvlprirkathvafckmknnyldatydvtvyegkddg	228

## RESULT 2

AY36728 standard; Protein; 352 AA.

AC AAY36728;

DT 27-SEP-1999 (first entry)

Human PGI<sub>2</sub> protein sequence.

KM PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;  
 KM cancer; prostate cancer; diagnosis; therapy; prostate specific antigen

05 Homo sapiens.

PN WO9932644-A2.

PD 01-JUL-1999.

22-DEC-1998; 98WO-IB02133.

PR 09-SEP-1998; 98US-0099658.

[illegible]

XX

XX

DR N-PSDB; AAZ00872.

PT	Use of a prostate
1	1
2	1
3	1
4	1
5	1
6	1
7	1
8	1
9	1
10	1
11	1
12	1
13	1
14	1
15	1
16	1
17	1
18	1
19	1
20	1
21	1
22	1
23	1
24	1
25	1
26	1
27	1
28	1
29	1
30	1
31	1
32	1
33	1
34	1
35	1
36	1
37	1
38	1
39	1
40	1
41	1
42	1
43	1
44	1
45	1
46	1
47	1
48	1
49	1
50	1
51	1
52	1
53	1
54	1
55	1
56	1
57	1
58	1
59	1
60	1
61	1
62	1
63	1
64	1
65	1
66	1
67	1
68	1
69	1
70	1
71	1
72	1
73	1
74	1
75	1
76	1
77	1
78	1
79	1
80	1
81	1
82	1
83	1
84	1
85	1
86	1
87	1
88	1
89	1
90	1
91	1
92	1
93	1
94	1
95	1
96	1
97	1
98	1
99	1
100	1

PT derived from it

PS Claim 7; Page 190; 385pp; English.

The invention relates to a mammalian Pgl gene and protein, and a set of Pgl biallelic markers. The Pgl polynucleotide and biallelic markers are used in a hybridisation assay, a sequencing assay, or in an allele-specific amplification assay for determining the identity of a nucleotide at a Pgl-related biallelic marker. The methods can be used to detect and to assess the risk of developing cancer or prostate cancer. Early-stage diagnosis of prostate cancer relies on prostate specific antigen (PSA) dosage. However, the effectiveness of this is limited due to its inability to discriminate between malignant and non-malignant affections of the organ. A need exists for both a reliable diagnostic procedure which would enable early-stage diagnosis, and for preventative and curative treatments of the disease. The Pgl gene can be used for detection of prostate cancer, and the risk of developing it in the future, and can also be used to determine therapies for the disease.

Sequence 352 AA;

Query Match	100.0%;	Score 228;	DB 20;	Length 352;
Best Local Similarity	100.0%;	Pred. No. 3.1e-222;		
Matches 228;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	1	MRYLPSVVLGGTATPYTLAAGCWRLLLSAFLPARFYQALMDRLKCYVQSMWLEFFFNENTG	60
Db	1	mtyllpsvvlvgtcapcltylaagvwrlisaflparfygaiddrlcyvqysmwlefffnetyg	60
Qy	61	VOILLPGDLPRKNENIITYLANHOSGVDMIVADILAIRONALGHNRYVLEKEGLKMLPLYGC	120
Db	61	vqlllygdlprknkenillylanhgstvdmivadillaingnalghnryvllkeglkmlplygc	120
Qy	121	YFAOHGIGTVKRSAKAFNEKEMRKNLQSYVDAGTPTMYLIFYEEGTRRYNEQTRKVLASQAF	180
Db	121	yfaohgigtvkrtsakafnekekrmknlqsyvdsagtpmylylffegtrrynepeqtrkvlasaqaf	180
Qy	181	AAQGLAVLKHVLPRIKATIHAFEDCKMKNITDAITDYTVYIEGKDDG	228
Db	181	aagglavlkhnvlprlkatihvafedckmknitdaydytvvyiegkddg	228

### RESULT 3

standard; Protein; 353 AA.

AC AAM39591;

DT	22-OCT-2001 (first entry)
----	---------------------------

DE	Human polypeptide	SEQ ID NO	2736.
22			

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;

Alzheimer's;  
KW

chemokinetic; t

XX :

XX  
XX

**XX :**

XX

XX

PR 25-APR-2000; 2000US-0552317.

PR 19-JUL-2000; 2000US-0620312.

PR 14-SEP-2000; 2000US-0662191.  
PR 19-OCT-2000; 2000US-0693036.  
PR 29-NOV-2000; 2000US-0727344.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
PI Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX  
XX WPI: 2001-442253/47.  
DR N-PSDB; AA158747.  
XX  
PT Novel nucleic acids and polypeptides, useful for treating disorders  
XX such as central nervous system injuries -  
XX  
XX Example 4; SEQ ID NO 2736; 10078bp; English.  
XX  
CC The invention relates to human nucleic acids (AA157798-AA161369) and  
CC the encoded polypeptides (AAM3642-AAM42213) with nootropic,  
CC immunosuppressant and cyrostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: immune system suppression,  
CC Actin/Inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.  
XX  
SQ Sequence 353 AA;  
XX

Query Match 100.0%; Score 228; DB 22; Length 353;  
Best Local Similarity 100.0%; Pred. No. 3.1e-222;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRLTSLVVLGTAFTYVLAAGWRLISAFLPARFOALDDRLCYQSWLFEFFENYTG 60  
DB 1 MYLLPSVVLGTAFTYVLAAGWRLISAFLPARFOALDDRLCYQSWLFEFFENYTG 60  
QY 61 VQILLYGDLPKKNENITLYLANHSTVDWIVADILAIRONALGHVRYLKEGLKMLPLYGC 120  
DB 61 VQILLYGDLPKKNENITLYLANHSTVDWIVADILAIRONALGHVRYLKEGLKMLPLYGC 120  
QY 121 YFAOHGIIYVRSKAKFNEKEMRNKLOSVDAGTPMVLVIFPESTRNPEQTKVLSAQAF 180  
DB 121 YFAOHGIIYVRSKAKFNEKEMRNKLOSVDAGTPMVLVIFPESTRNPEQTKVLSAQAF 180  
QY 181 AAOAGLAVLKHVLTPTKATHVAFDCMKNYLDAIVDTVYVEEKKDDGG 228  
DB 181 AAOAGLAVLKHVLTPTKATHVAFDCMKNYLDAIVDTVYVEEKKDDGG 228

RESULT 4  
AAB85532  
ID AAB85532 standard; protein; 353 AA.  
XX  
AC AAB85532;  
XX  
DE 25-SEP-2001 (first entry)  
XX  
XX Human secreted protein (clone id HLT084).  
XX  
XX Secreted protein; immunosuppressive; antiarthritic; antineumatic;  
KW antiproliferative; cytosolic; cardiac; vasotropic; cerebroprotective;  
KW nootropic; neuroprotective; antibacterial; virucide; fungicide; human;  
KW optalmallogical; gene therapy.

XX  
OS Homo sapiens.  
XX  
PN WO200155430-A1.  
XX  
XX 02-AUG-2001.  
PD  
PF 17-JAN-2001; 2001WO-US01431.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 12-SEP-2000; 2000US-0231968.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR, Olsen HS;  
PI Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH, Fiscella M;  
PI N J, Ruben SM, Barash SC;  
XX  
DR WPI: 2001-476220/51.  
DR N-PSDB; AAM46942.  
XX  
PS Claim 11; Page 450-451; 482pp; English.  
XX  
CC The invention provides novel human secreted proteins and polynucleotides  
CC encoding them. The secreted proteins can be expressed by standard  
CC recombinant methodology. The secreted proteins and polynucleotides are  
CC used to prevent, treat or ameliorate a medical condition in e.g. humans,  
CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They can  
CC also be used in diagnosing a pathological condition. The antibodies to  
CC the proteins can also be used in alleviating symptoms associated with the  
CC disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme  
CC linked immunosorbent assays (ELISA). Disorders which are diagnosed or  
CC treated include autoimmune diseases e.g. rheumatoid arthritis,  
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,  
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
CC e.g. cerebral ischemia, angiogenesis, nervous system disorders e.g.  
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and  
CC ocular disorders e.g. corneal infection. The polypeptides can also be  
CC used to aid wound healing and epithelial cell proliferation, to prevent  
CC skin aging due to sunburn, to maintain organs before transplantation, for  
CC supporting cell culture of primary tissues, to regenerate tissues and in  
CC chemotaxis. The polypeptides can also be used as a food additive or  
CC preservative to increase or decrease storage capabilities. The present  
CC sequence represents a human secreted protein.  
XX  
SQ Sequence 353 AA;  
XX

Query Match 100.0%; Score 228; DB 22; Length 353;  
Best Local Similarity 100.0%; Pred. No. 3.1e-222;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRLTSLVVLGTAFTYVLAAGWRLISAFLPARFOALDDRLCYQSWLFEFFENYTG 60  
DB 1 MYLLPSVVLGTAFTYVLAAGWRLISAFLPARFOALDDRLCYQSWLFEFFENYTG 60  
QY 61 VQILLYGDLPKKNENITLYLANHSTVDWIVADILAIRONALGHVRYLKEGLKMLPLYGC 120  
DB 61 VQILLYGDLPKKNENITLYLANHSTVDWIVADILAIRONALGHVRYLKEGLKMLPLYGC 120  
QY 121 YFAOHGIIYVRSKAKFNEKEMRNKLOSVDAGTPMVLVIFPESTRNPEQTKVLSAQAF 180  
DB 121 YFAOHGIIYVRSKAKFNEKEMRNKLOSVDAGTPMVLVIFPESTRNPEQTKVLSAQAF 180  
QY 181 AAOAGLAVLKHVLTPTKATHVAFDCMKNYLDAIVDTVYVEEKKDDGG 228  
DB 181 AAOAGLAVLKHVLTPTKATHVAFDCMKNYLDAIVDTVYVEEKKDDGG 228

## RESULT 5

AAB85552 standard; protein: 353 AA.

AAB85552:

25-SEP-2001 (first entry)

Human secreted protein (clone id HSL1A81).

Secreted protein: immunosuppressive; antiarthritic; antirheumatic;  
 antiapoptotic; cytoprotective; cardiatic; vasotropic; cerebroprotective;  
 neurotropic; neuroprotective; antibacterial; virucide; fungicide; human;  
 ophthalmological; gene therapy.

Homo sapiens.

MO200155430-A1.

02-AUG-2001.

17-JAN-2001; 2001WO-US01431.

31-JAN-2000; 2000US-0179065.

04-FEB-2000; 2000US-0180628.

12-SEP-2000; 2000US-0231968.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Komatsu G, Baker KP, Birse CE, Soppet DR, Olsen HS;

Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH, Fascella M;

N-J, Ruben SM, Barash SC;

WPI: 2001-476220/51.

N-PSDB: AAB46962.

17 Isolated nucleic acid molecules encoding human secreted proteins,

used to preventing, treating or ameliorating a medical condition

Claim 11; Page 465-466; 482pp; English.

The invention provides novel human secreted proteins and polynucleotides  
 encoding them. The secreted proteins can be expressed by standard  
 recombinant methodology. The secreted proteins and polynucleotides are  
 used to prevent, treat or ameliorate a medical condition in e.g. humans,  
 mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They can  
 also be used in diagnosing a pathological condition. The antibodies to  
 the proteins can also be used in alleviating symptoms associated with the  
 disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme  
 linked immunosorbent assays (ELISA). Disorders which are diagnosed or  
 treated include autoimmune diseases e.g. rheumatoid arthritis,  
 cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
 e.g. cerebral ischemia, angiogenesis, nervous system disorders e.g.  
 Alzheimer's disease, infections caused by bacteria, viruses and fungi and  
 other disorders e.g. corneal infection. The polypeptides can also be  
 used to aid wound healing and epithelial cell proliferation, to prevent  
 skin aging due to sunburn, to maintain organs before transplantation, for  
 supporting cell culture of primary tissues, to regenerate tissues and in  
 chemotaxis. The polypeptides can also be used as a food additive or  
 preservative to increase or decrease storage capabilities. The present  
 sequence represents a human secreted protein.

Sequence 353 AA;

Query Match 100.0%; Score 228; DB 22; Length 353;

Best Local Similarity 100.0%; Pred. No. 3.1e-222;

Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MRYLLPSVLLGTAPTYVLAAGWVRLSAPLPARFYQALDRLCYQSMVLEFFENYTG 60  
 1 MYLLPSVLLGTAPTYVLAAGWVRLSAPLPARFYQALDRLCYQSMVLEFFENYTG 60

QY 61 VQILLYGDLPKKNKIYLANHGSTVDWIVADILAIROMALGHVRYLKEGLKMLPYGC 120

Db 61 VQILLYGDLPKKNKIYLANHGSTVDWIVADILAIQNALGHVRYLKEGLKMLPYGC 120

QY 121 YFAOHGIGYVRSKAFNFKEMRNKLOSYVDAGTSMYLVIFEGRTYRNEQRYVLSAQAF 180

Db 121 YFAOHGIGYVRSKAFNFKEMRNKLOSYVDAGTSMYLVIFEGRTYRNEQRYVLSAQAF 180

QY 181 AAGRLAVLKHVLPRIKATHVAFDCMKKNYDAIYDTVYVEGKDDG 228

Db 181 AAGRLAVLKHVLPRIKATHVAFDCMKKNYDAIYDTVYVEGKDDG 228

## RESULT 6

AAV36729 standard; protein: 364 AA.

AAV36729;

27-SEP-1999 (first entry)

Human PGI protein sequence.

PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;

cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.

Homo sapiens.

W09932644-A2.

01-JUL-1999.

22-DEC-1998; 98WO-1B02133.

09-SEP-1998; 98US-0099658.

22-DEC-1997; 97US-0996306.

(GEST) GENSET.

Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;

WPI: 1999-405178/34.

N-PSDB: AA200870.

Use of a prostate cancer associated gene and biallelic markers  
 derived from it

Claim 7; Page 190-191; 385pp; English.

The invention relates to a mammalian PGI gene and protein, and a set of  
 PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
 used in a hybridisation assay, a sequencing assay, or in an  
 allele-specific amplification assay for determining the identity of a  
 nucleotide at a PGI-related biallelic marker. The methods can be used to  
 detect and to assess the risk of developing cancer or prostate cancer.  
 Early-stage diagnosis of prostate cancer relies on prostate specific  
 antigen (PSA) dosage. However, the effectiveness of this is limited due  
 to its inability to discriminate between malignant and non-malignant  
 affections of the organ. A need exists for both a reliable diagnostic  
 procedure which would enable early-stage diagnosis, and for preventative  
 and curative treatments of the disease. The PGI gene can be used for  
 detection of prostate cancer, and the risk of developing it in the  
 future, and can also be used to determine therapies for the disease.

Sequence 364 AA;

Query Match 100.0%; Score 228; DB 20; Length 364;

Best Local Similarity 100.0%; Pred. No. 3.2e-222;

Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MRYLLPSVLLGTAPTYVLAAGWVRLSAPLPARFYQALDRLCYQSMVLEFFENYTG 60  
 1 MRYLLPSVLLGTAPTYVLAAGWVRLSAPLPARFYQALDRLCYQSMVLEFFENYTG 60

Db	12	mryllpssvlllgtpcytylaagvrrlllsafipafiygaidrrlycyqsmvllffenytcg	71
Qy	61	VQILLGDPKRNKENIIYLANHGSTVDWIVADILAISQNALGHRVYLYKEGLKWLPLYGC	120
Db	72	vqlllygdlprkneniilylanhgstvdwlvadilaicqnalghvrvylkeglkwlpdygc	133
Qy	121	yFAOHGCIYVRSKAKFNENKEKRNKLQSYVDAGTPMYLYVPEEGTRIVPEQTKVLASQAF	186
Db	132	yfaohgciyvrskakfnekemrnlqsyvdagtpmylvifpegtrivpeqtkvlasaqaf	191
Qy	181	AAOGLAVLKHVLPRIKATFVAPADCKNKNYDAIYDVTVVYEGKDDGG	228
Db	192	aagqglavlkhvlprlkathvadcmknyldaiydvrvyegkddgg	239
RESULT 7			
ID	AAAM50128	standard; Protein; 364 AA.	
XX	AAAM50128;		
AC	AAAM50128;		
XX	21-DEC-2001 (first entry)		
DT	21-DEC-2001 (first entry)		
XX	Human acyltransferase 27417.		
DE	Human acyltransferase 27417.		
KW	Acyltransferase 27417; human; cell proliferation; cell migration;		
KW	cell differentiation; ovary cancer; brain cancer; colon cancer;		
KW	lung cancer; tumour; metastasis; sarcoma; carcinoma;		
KW	adenocarcinoma; antitumour; lipid metabolism; diagnosis; therapy.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FH	Peptide	1..49	
FT		/label= Signal_peptide	
FT	Protein	50..364	
FT		/label= Mature_protein	
FT	Domain	50..320	
FT		/label= N-terminal_domain	
FT	Domain	321..337	
FT		/label= Transmembrane_domain	
FT	Domain	338..364	
FT		/label= C-terminal_domain	
FT	Domain	71..363	
FT		/label= Acyltransferase_domain	
FT	Modified-site	68..71	
FT		/note= "Asn is N-glycosylated"	
FT	Modified-site	241..244	
FT		/note= "O-phosphorylated by cAMP- and cGMP-dependent protein kinase"	
FT	Modified-site	11..113	
FT		/note= "O-phosphorylated by protein kinase C"	
FT	Modified-site	144..146	
FT		/note= "O-phosphorylated by protein kinase C"	
FT	Modified-site	205..207	
FT		/note= "O-phosphorylated by protein kinase C"	
FT	Modified-site	317..319	
FT		/note= "O-phosphorylated by protein kinase C"	
FT	Modified-site	361..363	
FT		/note= "O-phosphorylated by protein kinase C"	
FT	Modified-site	95..98	
FT		/note= "O-phosphorylated by casein kinase II"	
FT	Modified-site	158..161	
FT		/note= "O-phosphorylated by casein kinase II"	
FT	Modified-site	246..249	
FT		/note= "O-phosphorylated by casein kinase II"	
FT	Modified-site	82..89	
FT		/note= "O-phosphorylated by tyrosine kinase"	
FT	Modified-site	219..226	
FT		/note= "O-phosphorylated by tyrosine kinase"	
FT	Modified-site	23..28	
FT		/note= "N-myristoylated"	

FT	Modified-site	130..135	/note= "N-myristoylated"
FT	Modified-site	330..335	/note= "N-myristoylated"
FT	Modified-site	352..357	/note= "N-myristoylated"
FT	Modified-site	230..343	/note= "N-myristoylated"
FT	Peptide	105..126	/note= "Amidation"
FT	Peptide	105..126	/note= "predicted leucine zipper"
PM	WO200173051-A2.		
XX			
PD	04-OCT-2001.		
XX			
PF	26-MAR-2001; 2001WO-US09633.		
XX			
PR	24-MAR-2000; 2000US-192092P.		
XX			
PA	(MILL-) MILLENNIUM PHARM INC.		
XX			
PI	Meyers RA, Rudolph-Owen L, Macbeth KJ, Tsai F, Williamson M;		
XX			
DR	WPI: 2001-626267/72.		
DR	N-PSDB; AAH27094.		
XX			
PT	New polynucleotide encodes a polypeptide for the diagnosis and		
PT	treatment of cancer and to alter cellular lipid metabolism comprises a		
PT	polynucleotide encoding human acyltransferases designated 46743 and		
PT	27417		
XX			
PS	Claim 9; Fig 8A-B; 150pp; English.		
XX			
CC	The present sequence is that of a novel human protein, termed		
CC	27417, which shows the structural characteristics of members of the		
CC	acyltransferase family, including an acyltransferase domain. The		
CC	invention provides novel acyltransferase 27417 nucleic acids and		
CC	polypeptides, as well as methods for detecting their presence, and		
CC	methods for screening for compounds that modulate their expression		
CC	or activity. Such compounds can be used to treat conditions		
CC	related to aberrant activity or expression of 27417 protein or		
CC	nucleic acid, such as conditions involving deficient cellular		
CC	proliferation, migration and/or differentiation. Also provided are		
CC	methods for inhibiting the proliferation or migration, or inducing		
CC	the killing, of a 27417-expressing cell, e.g. a hyperproliferative		
CC	and/or metastatic cell, by contact with a compound that modulates		
CC	the activity or expression of the 27417 protein or nucleic acid.		
CC	The 27417-expressing cell is found in the healthy or diseased		
CC	heart, blood vessels, kidney, skeletal muscle, brain or liver, or		
CC	especially in a solid tumour, a soft tissue tumour or a metastatic		
CC	lesion, a sarcoma, a carcinoma or an adenocarcinoma, and in		
CC	particular in a hyperproliferative and/or metastatic cell found in		
CC	ovarian, brain, colon or lung cancer. The compounds may also be		
CC	used to modulate lipid metabolism in a 27417-expressing cell.		
CC	Methods of disease diagnosis, e.g. by determining the		
CC	presence of a genetic alteration in a 27417 polypeptide, and for		
CC	evaluating the efficacy of a treatment of a disorder, are also		
CC	provided.		
XX			
XX			
SO	Sequence 364 AA:		
Query Match 100.0%; Score 228; DB 22; Length 364;			
Best Local Similarity 100.0%; Pred. No. 3 2e-222;			
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps			
QY	1 MKYLLPSVLLGTATPYVLAMGVWRLLSAFLPAPRYQALDLRLXCVYOSMVLFFENYTG 60		
Db	12 mtyllpsvlllgtatptylawgvwrllsaflparfygaldtrllcyovysmvlffennytg 71		
QY	61 VOILLGDLPPKKEIITILLAHQSTVDMIVADILAIRONALGHVRYVLKEGKKMLPLYGC 120		
Db	72 vqilllgdlppkknellilylanhgstvdwlvadilairqnalghvryvllkegllkwpilygc 131		

OY		121	VPAOHGCGIYVKKSAFNFENKEMRNKLQSYVDAGCPMWLVIFPEESTRRNPEDQTKVLSNQAF	180
DB		132	Yfagggggyvkvksaktnekernklqsyvdagtlpmylvlfpegtrnpegtkvljasgat	191
OY		181	AAORGLAVLKHLTPRIKATHVAFPDCMKRYVLDAIVTVTYVEEGRDDGG	228
DB		192	aagrglavlkhltpriktatvafdcmkryldaiydvtyvyegkdgg	239
	RESULT	8		
ID	AAU00665			
XX	AAU00665 standard; Pro-cln; 364 AA.			
AC	AAU00665;			
XX				
DT	07-SEP-2001 (first entry)			
XX				
DE	Human lysophosphatidic acid acyltransferase isoform LPAAT-epsilon.			
KW	Lysophosphatidic acid acyltransferase epsilon; LPAAT-epsilon; LPA; PA;			
KM	lysophosphatidic acid; phosphatidic acid; acylation; cellular activation;			
RN	phospholipid signalling; mitogenesis; inflammation; autoimmune disease;			
KX	oncology; cancer; obesity; gene therapy.			
XX				
OS	Homo sapiens.			
PN	WO200134782-A1.			
PD	17-MAY-2001.			
XX				
PF	02-NOV-2000; 2000MO-US30193.			
PR	09-NOV-1999; 99US-0436319.			
XX				
PA	(CELL-) CELL THERAPEUTICS INC.			
XX				
P1	Leung DW;			
DR	WPI: 2001-335920/35.			
DR	N-PSTB; AAS00649.			
XX				
PT	Novel isolated human isoform of lysophosphatidic acid			
PT	acyltransferase-epsilon useful for diagnostic, therapeutic and			
PT	screening purposes -			
XX				
PS	Clam 1; Fig 1; 48pp; English.			
CC	The sequence represents a human lysophosphatidic acid acyltransferase			
CC	(LPAAT) isoform, LPAAT-epsilon. LPAAT catalyses the acylation of			
CC	lysophosphatidic acid (LPA) to phosphatidic acid (PA). LPA and PA have			
CC	been identified as phospholipid signalling molecules that affect a wide			
CC	range of biological responses. PA is involved in cellular activation and			
CC	mitogenesis. Compounds that block PA generation and hence diminish lipid			
CC	biosynthesis and the signal involved in cell activation are of			
CC	therapeutic interest in the areas of inflammation and oncology (e.g.			
CC	autoimmune diseases and cancer) as well as obesity treatment.			
CC	LPAAT-epsilon and its corresponding DNA can be used in screening assays			
CC	to detect agents that stimulate or inhibit the activity of LPAAT and,			
CC	therefore, PA. The DNA is useful in tests to detect the presence or			
CC	expression of LPAAT-epsilon in relation to certain diseases and			
CC	conditions, and in disease prevention and treatment. The sequences of the			
CC	invention are also useful for diagnosis of diseases and conditions in			
CC	which the expression of LPAAT enzyme is abnormal.			
XX				
SQ	Sequence 364 AA;			
	Query Match 100.0%; Score 228; DB 22; Length 364;			
	Best Local Similarity 100.0%; Pred. No. 3,2e-222;			
	Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0			
OY	1 KRYLLPVSULLCTAPTYVLAWGVWRLLSALPARFYQALDRLCYVGSMVLEFFENYTG	60		

```

Db      12 myrlllpsvlllslaptcyvlawgvrllseflparfyrqaldcllycvgsamvllfflenyvg 71
        |||
Qy      61 VOILLYGDLPPKKNENTITVIANHQSIVDWIVADILAIRQNALCHVRVYLKEGLKMLPLTGC 120
        |||
Db      72 vqlllygdlpknkhenillylanhqstcdwlvadllafrqnalghvryvllkeglkwpilpygc 131
        |||
Qy      121 YFAOHGCIYVKSASFKNEMKNEKEMRNKLOSIVDADCTPMWLVFEPGSTRNPEQRTVLSASQAF 180
        |||
Db      132 ylaqnggilyvksakfnekeemrnlqsyvdagctpmlylvlfpegrlrynpqclvksasqaf 191
        |||
Qy      181 AAOGLAVLKLHVLTPEIKATHVAFDCKMKNYLDALIVDTVTVYEGKDDCG 228
        |||
Db      192 aagrglavllkhvltprlkathvafdomknyldalivdtvvyegkddgg 239
        |||

RESULT 9
AAV36753
ID      AAV36753 standard; Protein; 185 AA.
XX
XX      AAV36753;
AC
XX
XX      27-SEP-1999 (first entry)
DE
XX      PGI splice variant allele protein sequence.
DE
XX      PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;
XX      cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.
XX
XX      Homo sapiens.
OS
XX
XX      WO9932644-A2.
XX
XX      01-JUL-1999.
XX
XX      22-DEC-1998; 98WO-1B02133.
XX
XX      09-SEP-1998; 98US-0099658.
XX      22-DEC-1997; 97US-0096306.
XX
XX      (GEST ) GENSET.
XX
XX      Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;
PI
XX      WPT; 1999-405178/34.
XX      DR
XX      N-PSDB; AAZ00979.
XX
XX      Use of a prostate cancer associated gene and biallelic markers
XX      derived from it
XX
XX      Claim 7; Page 255; 385pp; English.
XX
XX      The invention relates to a mammalian PGI gene and protein, and a set of
XX      PGI biallelic markers. The PGI polynucleotide and biallelic markers are
XX      used in a hybridisation assay, a sequencing assay, or in an
XX      allele-specific amplification assay for determining the identity of a
XX      nucleotide at a PGI-related biallelic marker. The methods can be used to
XX      detect and to assess the risk of developing cancer or prostate cancer.
XX      Early-stage diagnosis of prostate cancer relies on prostate specific
XX      antigen (PSA) dosage. However, the effectiveness of this is limited due
XX      to its inability to discriminate between malignant and non-malignant
XX      affections of the organ. A need exists for both a reliable diagnostic
XX      procedure which would enable early-stage diagnosis, and for preventative
XX      and curative treatments of the disease. The PGI gene can be used for
XX      detection of prostate cancer, and the risk of developing it in the
XX      future, and can also be used to determine therapies for the disease.
XX
XX      Sequence 185 AA;
XX
Query Match      81.1%; Score 185; DB 20; Length 185;
Best Local Similarity 100.0%; Pred. No. 5.6e-179;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Best Local Similarity 100.0%; Pred. No. 9.2e-179;  
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MRLLPSVILGTAPTVILAMGVARLLSAPLPAFFYQALDDRLXCVCQSNVLEFFENYTG	60
Db	1	mrtyllpsvllgtapcylvllagvwrltsatlpafiygaldrllycvysmwllffenytg	60
QY	61	VOILLYGDLPKNKENIITYLANHOSTVDMIVADILAIQNALIGHVRYLKEGLKWLPLYGC	120
Db	61	vglllygdrlprknkennllylanhgstvcdwlvadlalatqmalghnrvylkeglkwlpriygc	120
QY	121	YFAOHGGLIYKRSKKEFEKEKRNKLQSYVDAQIPMYIVIREGSTRVPEQTKVLASQAR	180
Db	121	yfaahgg9llykrsakfnekemrnklqsyvdagfpmlyivlregrlrvpeqctkylasqarf	180
QY	181	AAORG 185	
Db	181	aagrg 185	

RESULT	10
AAV36751	
ID	AAV36751 standard; Protein; 315 AA

AC	AAV36751;
XX	
DT	27-SEP-1999 (first entry)
...	

PG1 splice variant allele protein sequence.

KM PGI gene; biallelic marker; human; PGI-related biallelic marker;  
KM cancer; prostate cancer; diagnosis; therapy; prostate specific antigen

Homo sapiens.

PN W09932644-A2.

PD 01-JUL-1999.

PF 22-DEC-1998; 98WO-IB02133.

PR 09-SEP-1998; 98US-0099658.

XX  
EN 24 DEC 1951, 3105-0330300.  
XX

XX  
XX  
(0001) GENSEL  
XX  
XX

XX  
P1  
Blumenfeld M, Bouguetere L, Chumakov I, Cohen D;

DR WFL; 1993-403178/34.  
DR N-PSDB; AAZ00977.

XX PT Use of a prostate cancer associated gene and biallelic markers

PT derived from it  
XX  
PS Claim 7; Page 253-254; 385pp; English.

CC The invention relates to a mammalian Pgl gene and protein, and a set of  
CC Pgl biallelic markers. The Pgl polynucleotide and biallelic markers are  
CC used in a hybridisation assay, a sequencing assay, or in an  
CC allele-specific amplification assay for determining the identity of a  
CC nucleotide at a Pgl-related biallelic marker. The methods can be used to  
CC detect and to assess the risk of developing cancer or prostate cancer.  
CC Early-stage diagnosis of prostate cancer relies on prostate specific  
CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
CC to its inability to discriminate between malignant and non-malignant  
CC affections of the organ. A need exists for both a reliable diagnostic  
CC procedure which would enable early-stage diagnosis, and for preventative  
CC and curative treatments of the disease. The Pgl gene can be used for  
CC detection of prostate cancer, and the risk of developing it in the  
CC future, and can also be used to determine therapies for the disease.

Query Match	81.1%;	Score 185;	DB 20;	Length 315;
-------------	--------	------------	--------	-------------

	Best Local Similarity	100.0%	Pred.	No. 9, 2e-179					
	Matches	185	Conservative	0	Mismatches	0	Indels	0	Gaps
QY	1	MYLLPSVVLGTAATYYLAWGWRFLSAFLPARFYQALDRLKCYQSNVLEFFFEYTYG	60						
Db	1	MYLLPSVVLGTAATYYLAWGWRFLSAFLPARFYQALDRLKCYQSNVLEFFFEYTYG	60						
QY	61	VOILLYGDLPKKRENIITLANQSTQVDMIVADIIAIRONALGHRVYLKGLKMLPLYGC	120						
Db	61	VQILLYGDLPKKEIIITLANQSTQVDMIVADIIAIRNALGHRVYLKGLKMLPLYGC	120						
QY	121	YFNOHGCIYVRKSAFENKEMKMKIQTQSVVDAGTGMVYLVIPEEGRVRYNDEQTKVLSASQAF	180						
Db	121	YFAGNQGIVYKSAFENKEMKMKIQSYVDAGTGMVYLVIPEGRITRYNDEQTKVLSASQAF	180						
QY	181	AAORG	185						
Db	181	AAQRG	185						

RESULT 11

ID	AAV36752	standard; Protein; 300 AA
XX		
AC	AAV36752;	

DT 27-SEP-1999 (first entry)

PG1 splice variant allele protein sequence.

KW cancer; prostate

AA OS Homo sapiens

PN W09932644-A2

01-JUL-1999.

AA 22-DEC-1998; 98WO-IB02133.  
PF

XX 09-SEP-1998; 98US-0099658.  
PR

44-DEC-1991;  
PR  
XX

PA (GEST) GENSET.  
XX  
XX

PI Blumenfeld M, Bougu  
XX

DR WPT; 1999-405178/34.  
DR N-PSDB: AAZ00978.  
DR

XX	Use of a prostate cancer associated gene derived from it
PT	

PS Claim 7, page 254-255; 385pp; English.

Sequence 300 AA;







SQ Sequence 269 AA:

Query Match 63.2%; Score 144; DB 22; Length 269;  
 Best Local Similarity 100.0%; Pred. No. 2,4e-137;  
 Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 TVDMIVADILAIROMALGHVRYLKEGLKWLPLYGCYFAOHGCIYVRSRKFNEKEMRNK 144  
 |||  
 Db 1 tvdmivadiqlatrqnalghvrylkeglkwlplygcylfaohgciyvrksakfnekeemrnk 60

QY 145 LOSYVADGTPMYLVIPEEGTRVNEPCTKVLSSQAFPAOAGLAVLKIVLPRIATIVAF 204  
 |||  
 Db 61 lsgyvdaqgtpmylviipeegtrvnepctkvlssaqaafagrlavlkivlprikathvaf 120

QY 205 DCMKNYLDALYDVYVTEGKDDCG 228  
 |||  
 Db 121 dcmknyldalydvrvtyegkddgg 144

RESULT 14  
 AAB93595  
 ID AAB93595 standard; Protein; 353 AA.  
 XX  
 AC AAB93595;  
 XX  
 DT 26-JUN-2001 (first entry)  
 XX  
 DE Human protein sequence SEQ ID NO:13028.  
 XX  
 KM Human; primer: detection; diagnosis; antisense therapy; gene therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP1074617-AZ.  
 XX  
 PD 07-FEB-2001.  
 XX  
 PF 28-JUL-2000; 2000EP-0116126.  
 XX  
 PR 29-JUL-1999; 99JP-0248036.  
 PR 27-AUG-1999; 99JP-0300253.  
 PR 11-JAN-2000; 2000JP-0118776.  
 PR 02-MAY-2000; 2000JP-0183767.  
 PR 09-JUN-2000; 2000JP-0241899.  
 XX  
 PA (HELI-) HELIX RES INST.  
 XX  
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX  
 DR WPI; 2001-318749/34.  
 XX  
 PT Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs -  
 XX  
 PS Claim 8; SEQ ID 13028; :537pp + CD ROM; English.  
 XX  
 CC The present invention describes primer sets for synthesizing 5602  
 CC full-length cDNAs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and

CC In gene therapy. The primers are useful for synthesizing polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialized methods. AAH03166 to AAH13628 and  
 CC AAH1633 to AAH18742 represent human cDNA sequences; AAB92446 to  
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH1632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.  
 CC  
 SQ Sequence 353 AA:

Query Match 63.2%; Score 144; DB 22; Length 353;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-137;  
 Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVVLGTRPTVYVLAAGVWRLLSATLPARFYALDDRLCYQVSNVLEFFENYTG 60  
 |||  
 Db 1 mryllpsvvlgtaptvylawgyvrrllsatlparyaldrlcyvqsnvleffenytg 60

QY 61 VOILLYGDLPKKNENIITYLANHOSTVDMIVADILAIROMALGHVRYLKEGLKWLPLYGC 120  
 |||  
 Db 61 vqillygdlpkkneniitylanhostvdmivadiqlatrqnalghvrylkeglkwlplygc 120

QY 121 YFAOHGCIYVRSRKFNEKEMRNK 144  
 |||  
 Db 121 yfaohgciyvrsakfnekeemrnk 144

RESULT 15  
 AAM41377  
 ID AAM41377 standard; Protein; 372 AA.  
 XX  
 AC AAM41377;  
 XX  
 DT 22-OCT-2001 (first entry)  
 XX  
 DE Human polypeptide SEQ ID NO 6308.  
 XX  
 KM Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KM peripheral nervous system; neuropathy; central nervous system; CNS;  
 KM Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KM chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KM leukaemia.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200153312-A1.  
 XX  
 PD 26-JUL-2001.  
 XX  
 PF 26-DEC-2000; 2000MO-US34263.  
 XX  
 PR 21-JAN-2000; 2000US-0488725.  
 PR 25-APR-2000; 2000US-0552317.  
 PR 09-JUL-2000; 2000US-0598042.  
 PR 19-JUL-2000; 2000US-0620312.  
 PR 03-AUG-2000; 2000US-0653450.  
 PR 14-SEP-2000; 2000US-0662191.  
 PR 19-OCT-2000; 2000US-0693036.  
 PR 29-NOV-2000; 2000US-0727344.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
 XX  
 DR WPI: 2001-442253/47.  
 DR N-PDB; AAI60533.  
 XX

PT Novel nucleic acids and polypeptides, useful for treating disorders  
 PT such as central nervous system injuries -

XX  
 PS Example 2; SEQ ID NO 6308; 10078bp; English.

CC The invention relates to human nucleic acids (AA157798-AA161369) and  
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemias and  
 CC C.N.S disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.

CC  
 SQ Sequence 372 AA:

Query Match 55.7%; Score 127; DB 22; Length 372;  
 Best Local Similarity 99.6%; Pred No. 5.2e-120;

Matches 227; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MRYLLPSVVLGTAPTYVLANGVWRLLSAFLPARFYQALDRLCYVYQSMVLEFFENYTG 60  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 20 MYLLPSVVLGTAPTYVLANGVWRLLSAFLPARFYQALDRLCYVYQSMVLEFFENYTG 79  
 OY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 61 VOILLXGDLPRKKNENIYIANHQSIVDWIVADILAIRONALGHVRYVLEKGLKMLPLCYC 120  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 80 VQILLIYGDLPKKNENIYIANHQSIVDWIVADILAIRONALGHVRYVLEKGLKMLPLCYC 139  
 OY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 121 YFAOHGAGIVKRSKAKFNEKEMRNKLOSIVDAGTPMVLTFPEGSTRYNPEOTKVLSSQAF 180  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 140 YFAOHGAGIVKRSKAKFNEKEMRNKLOSIVDAGTPMVLTFPEGSTRYNPEOTKVLSSQAF 199  
 OY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 181 AAORGLAVLKHLVTPRIKATHVAFDCKMKNYLDIYDVTVVYEGKDDGG 228  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 200 AAGRGLAVLKHLVTPRIKATHVAFDCKMKNYLDIYDVTVVYEGKDDGG 247

Search completed: August 28, 2002, 11:18:45  
 Job time: 583 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:19:21 ; Search time 30.82 Seconds  
(without alignments)  
180.696 Million cell updates/sec

Title: US-09-853-526-70

Perfect score: 228  
Sequence: 1 MRYLLPSVLLGTAPTYVLA.....NYLDAIVDVTVYEGKDDG 228

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 231628 seqs, 24425594 residues

Word size : 0

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents-AA:\*

- 1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*
- 2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*
- 3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*
- 4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*
- 5: /cgn2\_6/ptodata/2/1aa/PTCUS.COMB.pep:\*
- 6: /cgn2\_6/ptodata/2/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	228	100.0	228	4	US-09-338-907-70 Sequence 70, App1
2	228	100.0	228	4	US-09-218-207-70 Sequence 70, App1
3	228	100.0	353	2	US-08-996-306-4 Sequence 4, App1
4	228	100.0	353	4	US-09-338-907-4 Sequence 4, App1
5	228	100.0	353	4	US-09-218-207-4 Sequence 4, App1
6	228	100.0	364	2	US-08-996-306-5 Sequence 5, App1
7	228	100.0	364	4	US-09-338-907-5 Sequence 5, App1
8	228	100.0	364	4	US-09-218-207-5 Sequence 5, App1
9	185	81.1	185	4	US-09-338-907-136 Sequence 136, App
10	185	81.1	185	4	US-09-218-207-136 Sequence 136, App
11	185	81.1	315	4	US-09-338-907-134 Sequence 134, App
12	185	81.1	315	4	US-09-218-207-134 Sequence 134, App
13	184	80.7	300	4	US-09-338-907-135 Sequence 135, App
14	184	80.7	300	4	US-09-218-207-135 Sequence 135, App
15	154	67.5	182	4	US-09-338-907-133 Sequence 133, App
16	154	67.5	182	4	US-09-218-207-133 Sequence 133, App
17	105	46.1	291	4	US-09-338-907-127 Sequence 127, App
18	105	46.1	291	4	US-09-218-207-127 Sequence 127, App
19	85	37.3	97	4	US-09-338-907-132 Sequence 132, App
20	85	37.3	97	4	US-09-218-207-132 Sequence 132, App
21	74	32.5	261	4	US-09-338-907-128 Sequence 128, App
22	74	32.5	261	4	US-09-218-207-128 Sequence 128, App
23	62	27.2	66	4	US-09-338-907-131 Sequence 131, App
24	62	27.2	66	4	US-09-218-207-131 Sequence 131, App
25	62	27.2	68	4	US-09-338-907-130 Sequence 130, App
26	62	27.2	68	4	US-09-218-207-130 Sequence 130, App
27	62	27.2	77	4	US-09-338-907-125 Sequence 125, App

28	62	27.2	77	4	US-09-218-207-125	Sequence 125, App
29	62	27.2	90	4	US-09-338-907-129	Sequence 129, App
30	62	27.2	90	4	US-09-218-207-129	Sequence 129, App
31	62	27.2	238	4	US-09-338-907-126	Sequence 126, App
32	62	27.2	238	4	US-09-218-207-126	Sequence 126, App
33	62	14.0	354	4	US-09-338-907-74	Sequence 74, App1
34	32	14.0	354	4	US-09-218-207-74	Sequence 74, App1
35	9	3.9	9	2	US-08-996-306-15	Sequence 15, App1
36	9	3.9	9	4	US-09-338-907-15	Sequence 15, App1
37	9	3.9	9	4	US-09-218-207-15	Sequence 15, App1
38	8	3.5	9	2	US-08-996-306-18	Sequence 18, App1
39	8	3.5	9	4	US-09-338-907-18	Sequence 18, App1
40	8	3.5	9	4	US-09-218-207-18	Sequence 18, App1
41	7	3.1	12	2	US-08-224-625-1	Sequence 1, App1
42	7	3.1	12	2	US-08-254-404-1	Sequence 1, App1
43	7	3.1	12	2	US-08-327-451E-1	Sequence 1, App1
44	7	3.1	12	2	US-08-458-109-1	Sequence 1, App1
45	7	3.1	12	3	US-08-231-196-1	Sequence 1, App1

ALIGNMENTS

RESULT 1  
US-09-338-907-70  
; Sequence 70, Application US/09338907  
; Patent No. 6265546  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marla  
; APPLICANT: Ilya, Chumakov  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: PROSTATE CANCER GENE  
; FILE REFERENCE: GENSET.18CP1CP  
; CURRENT APPLICATION NUMBER: US/09/338, 907  
; EARLIER FILING DATE: 1999-06-23  
; EARLIER APPLICATION NUMBER: 08/996, 306  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: 60/099, 658  
; EARLIER FILING DATE: 1998-09-09  
; EARLIER APPLICATION NUMBER: 09/218, 207  
; EARLIER FILING DATE: 1998-12-22  
; NUMBER OF SEQ ID NOS: 578  
; SOFTWARE: Patent.pm  
; SEQ ID NO 70  
; LENGTH: 228  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-338-907-70

Query Match 100.0%; Score 228; DB 4; Length 228;  
Best Local Similarity 100.0%; Pred. No. 1.1e-221; Indels 0; Gaps 0;  
Matches 228; Conservative 0; Mismatches 0;

QY	1	MRYLLPSVLLGTAPTYVLAWGVRRLLSAFLPAREYQALDRLCYQSMVLFFENYTG	60
Db	1	MRYLLPSVLLGTAPTYVLAWGVRRLLSAFLPAREYQALDRLCYQSMVLFFENYTG	60
QY	61	VQILLXGDLPRKNENIYLANHSTVDMIVADILAIKRONALGHVRYLKGKMLPLYGC	120
Db	61	VQILLXGDLPRKNENIYLANHSTVDMIVADILAIKRONALGHVRYLKGKMLPLYGC	120
QY	121	YFAOHGSIYKRSKAKFNEKMRKKLOSVDAGTIPMYLVIPESTRNPEOTKVLASQAF	180
Db	121	YFAOHGSIYKRSKAKFNEKMRKKLOSVDAGTIPMYLVIPESTRNPEOTKVLASQAF	180
QY	181	AAORGIAVLKHLVLPRIKATHVAFDCKKNYLDIYDTVYVEGKDDG	228
Db	181	AAORGIAVLKHLVLPRIKATHVAFDCKKNYLDIYDTVYVEGKDDG	228
RESULT	2		

US-09-218-207-70  
; Sequence 70, Application US/09218207  
; Patent No. 6346381  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Ilya, Chumakov  
; TITLE OF INVENTION: Prostate cancer gene  
; FILE REFERENCE: GENSET.018CPI  
; CURRENT APPLICATION NUMBER: US/09/218, 207  
; EARLIER FILING DATE: 1993-12-22  
; EARLIER APPLICATION NUMBER: 08/996, 306  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: 60/099, 658  
; EARLIER FILING DATE: 1996-09-09  
; NUMBER OF SEQ ID NOS: 576  
; SOFTWARE: Patent.pm  
; SEQ ID NO 70  
; LENGTH: 228  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-218-207-70

Query Match 100.0%; Score 228; DB 4; Length 228;  
Best Local Similarity 100.0%; Pred. No. 1.1e-221;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTAPTYVLLANGVWRLLSAPLPARYQALDRLKCYVSMVLPFFENTG 60  
DB 1 MRYLLPSVLLGTAPTYVLLANGVWRLLSAPLPARYQALDRLKCYVSMVLPFFENTG 60  
QY 61 VQILLYGDLPRKNENIITLHNMOSTVDIVADILAIRONALGHVRYVKEGLKMLPL/GC 120  
DB 61 VQILLYGDLPRKNENIITLHNMOSTVDIVADILAIRONALGHVRYVKEGLKMLPL/GC 120  
QY 121 YFAOHGCIYVRSKAFENKEKRNKLSQSYVDAGTPMYLVIPEGTRYNDEQTRVLSASQAF 180  
DB 121 YFAOHGCIYVRSKAFENKEKRNKLSQSYVDAGTPMYLVIPEGTRYNDEQTRVLSASQAF 180  
QY 181 AAQGLAVLKHVLPRIKATHVAFDCKMKNYLDATYDVTVYEGKDDGG 228  
DB 181 AAQGLAVLKHVLPRIKATHVAFDCKMKNYLDATYDVTVYEGKDDGG 228

RESULT 3  
US-08-996-306-4  
; Sequence 4, Application US/08996306  
; Patent No. 5945522  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Chumakov, Ilya  
; APPLICANT: Blumenfeld, Marta  
; TITLE OF INVENTION: Prostate cancer gene  
; NUMBER OF SEQUENCES: 68  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson & Bear  
; STREET: 501 West Broadway  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92101-3505  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; OPERATING SYSTEM: WIN95  
; SOFTWARE: Word  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/996, 306  
; FILING DATE:  
; CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:  
NAME: Israelson, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: GENSET.018A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 235-8550  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 353 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: SINGLE  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PROTEIN  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 1..33  
IDENTIFICATION METHOD: Rao and Argos method  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 4..20  
IDENTIFICATION METHOD: Klein, Kanehisa and DeLisi method  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 4..24  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony and Wall method  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 12  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 50..70  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony and Wall method  
FEATURE:  
NAME/KEY: potential N-glycosylation site  
LOCATION: 57  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 76..96  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony and Wall method  
FEATURE:  
NAME/KEY: potential Tyrosine kinase phosphorylation site  
LOCATION: 78  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Casein kinase II phosphorylation site  
LOCATION: 84  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Leucine zipper pattern  
LOCATION: 94..115  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 119  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 133  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Casein kinase II phosphorylation site  
LOCATION: 147  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 194  
IDENTIFICATION METHOD: prosite match

```
FEATURE:
NAME/KEY: potential Tyrosine kinase phosphorylation site
LOCATION: 215
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Tyrosine sulfatation site
LOCATION: 221
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential cAMP- and cGMP-dependent protein kinase phosphorylation site
LOCATION: 223
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Casein kinase II phosphorylation site
LOCATION: 225
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Protein kinase C phosphorylation site
LOCATION: 306
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Transmembrane helix
LOCATION: 310..330
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony and Wall method
FEATURE:
NAME/KEY: potential N-myristoylation site
LOCATION: 319
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential N-myristoylation site
LOCATION: 323
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Amidation site
LOCATION: 329
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Transmembrane helix
LOCATION: 333..353
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony and Wall method
FEATURE:
NAME/KEY: potential N-myristoylation site
LOCATION: 341
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Protein kinase C phosphorylation site
LOCATION: 350
IDENTIFICATION METHOD: prosite match
US-08-996-306-4
```

Query Match 100.0%; Score 228; DB 2; Length 353;  
Best Local Similarity 100.0%; Pred. No. 1.6e-221;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 MRLLEPSVLLGAPPTVYVAMGWRLISAFLEPARFYQALDDRLCYQVQSQVLEFFENYTG 60
DB 1 MRLLEPSVLLGAPPTVYVAMGWRLISAFLEPARFYQALDDRLCYQVQSQVLEFFENYTG 60
OY 61 VOILLXGDLPRKNKENTIIYLANHOSFVDWIVADILAIKRONALGHVRYVLEKGLKMLPLYGC 120
DB 61 VOILLXGDLPRKNKENTIIYLANHOSFVDWIVADILAIKRONALGHVRYVLEKGLKMLPLYGC 120
OY 121 YFAOHGIGIYKRSAKNEKEMRKLOSVDAGTPTWLVIFPESTRINPEQTKVLSAQAF 180
DB 121 YFAOHGIGIYKRSAKNEKEMRKLOSVDAGTPTWLVIFPESTRINPEQTKVLSAQAF 180
OY 181 AAORGIAVLKHYLTPRIKATHVAFDCMKNYLDAIVDTVYVEKDGOG 228
DB 181 AAORGIAVLKHYLTPRIKATHVAFDCMKNYLDAIVDTVYVEKDGOG 228
```

RESULT 4

```
US-09-338-907-4
; Sequence 4, Application US/09338907
; Patent No. 6265546
GENERAL INFORMATION:
APPLICANT: Cohen, Daniel
APPLICANT: Blumenfeld, Marta
APPLICANT: Ilya, Chumakov
APPLICANT: Bouguenere, Lydie
TITLE OF INVENTION: PROSTATE CANCER GENE
FILE REFERENCE: GENSET.18CPICP
CURRENT APPLICATION NUMBER: US/09/338,907
CURRENT FILING DATE: 1999-06-23
EARLIER APPLICATION NUMBER: 08/996,306
EARLIER FILING DATE: 1997-12-22
EARLIER APPLICATION NUMBER: 60/099,658
EARLIER FILING DATE: 1998-09-09
EARLIER APPLICATION NUMBER: 09/218,207
EARLIER FILING DATE: 1998-12-22
NUMBER OF SEQ ID NOS: 578
SOFTWARE: Patent.pm
SEQ ID NO 4
LENGTH: 353
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: HELIX
LOCATION: 1..33
OTHER INFORMATION: Rao and Argos identification method, potential helix
FEATURE:
NAME/KEY: HELIX
LOCATION: 4..20
OTHER INFORMATION: Klein, Kanehisa and Delisi identification method, potent
FEATURE:
NAME/KEY: MYRISTATE
LOCATION: 12..16
OTHER INFORMATION: Prosite match
FEATURE:
NAME/KEY: HELIX
LOCATION: 50..70
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
FEATURE:
NAME/KEY: CARBOHYD
LOCATION: 57..59
OTHER INFORMATION: Prosite match
FEATURE:
NAME/KEY: HELIX
LOCATION: 76..96
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
FEATURE:
NAME/KEY: PHOSPHORYLATION
LOCATION: 78
OTHER INFORMATION: potential Tyrosine kinase site, Prosite match
FEATURE:
NAME/KEY: PHOSPHORYLATION
LOCATION: 84
OTHER INFORMATION: potential caseine kinase II site, Prosite match
FEATURE:
NAME/KEY: SITE
LOCATION: 94..115
OTHER INFORMATION: potential Leucine zipper site, Prosite match
FEATURE:
NAME/KEY: MYRISTATE
LOCATION: 119..123
OTHER INFORMATION: potential site, Prosite match
FEATURE:
NAME/KEY: PHOSPHORYLATION
LOCATION: 133
OTHER INFORMATION: potential protein kinase C, Prosite match
FEATURE:
```

```

: NAME/KEY: PHOSPHORYLATION
: LOCATION: 147
: OTHER INFORMATION: potential caseine kinase II site, Prosite match
: FEATURE:
: NAME/KEY: PHOSPHORYLATION
: LOCATION: 194
: OTHER INFORMATION: potential protein kinase C, Prosite match
: FEATURE:
: NAME/KEY: PHOSPHORYLATION
: LOCATION: 215
: OTHER INFORMATION: potential Tyrosine kinase site, Prosite match
: FEATURE:
: NAME/KEY: SULFATATION
: LOCATION: 221
: OTHER INFORMATION: Prosite match
: FEATURE:
: NAME/KEY: PHOSPHORYLATION
: LOCATION: 233
: OTHER INFORMATION: potential cAMP and cGMP dependant protein kinase site, Prosite match
: FEATURE:
: NAME/KEY: PHOSPHORYLATION
: LOCATION: 235
: OTHER INFORMATION: potential caseine kinase II site, Prosite match
: FEATURE:
: NAME/KEY: PHOSPHORYLATION
: LOCATION: 306
: OTHER INFORMATION: potential protein kinase C, Prosite match
: FEATURE:
: NAME/KEY: HELIX
: LOCATION: 310..330
: OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, Prosite match
: FEATURE:
: NAME/KEY: MYRISTATE
: LOCATION: 319..323
: OTHER INFORMATION: Prosite match
: FEATURE:
: NAME/KEY: MYRISTATE
: LOCATION: 323..327
: OTHER INFORMATION: Prosite match
: FEATURE:
: NAME/KEY: AMIDATION
: LOCATION: 329
: OTHER INFORMATION: Prosite match
: FEATURE:
: NAME/KEY: HELIX
: LOCATION: 333..353
: OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, Prosite match
: FEATURE:
: NAME/KEY: MYRISTATE
: LOCATION: 341..345
: OTHER INFORMATION: Prosite match
: FEATURE:
: NAME/KEY: PHOSPHORYLATION
: LOCATION: 350
: OTHER INFORMATION: potential protein kinase C, Prosite match
US-09-338-907-4
```

Query Match 100.0%; Score 228; DB 4; Length 353;

Best Local Similarity 100.0%; Pred. No. 1.6e-221; Indels 0; Gaps 0;

Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 MRYLLPSVLLGTAFTYVLAHWRLSAFLPARFYQALDRLCYVQSMVLEFFENTY 60
Db 1 MRYLLPSVLLGTAFTYVLAHWRLSAFLPARFYQALDRLCYVQSMVLEFFENTY 60
Qy 61 VOILLYDLPKKNENIYYLANHSTVDWIVADILAIRONALGHVRYVLEGLKMLPLXGC 120
Db 61 VOILLYDLPKKNENIYYLANHSTVDWIVADILAIRONALGHVRYVLEGLKMLPLXGC 120
Qy 121 YFAOHGGLIYVRSKAKFNEKEMRKLQSVVDAGTSMYLVIFPEGTRRYNPEQTKVLSASQAF 180
Db 121 YFAOHGGLIYVRSKAKFNEKEMRKLQSVVDAGTSMYLVIFPEGTRRYNPEQTKVLSASQAF 180
```

```

Qy 181 AAGRLAVLKHVLPRIKATHVAFDCMKNYLDAIYDVTVYEGKDDG 228
Db 181 AAGRLAVLKHVLPRIKATHVAFDCMKNYLDAIYDVTVYEGKDDG 228
```

```

RESULT 5
US-09-218-207-4
: Sequence 4, Application US/09218207
: Patent No. 6346381
: GENERAL INFORMATION:
: APPLICANT: Cohen, Daniel
: APPLICANT: Blumenfeld, Maria
: APPLICANT: Ilya, Chumakov
: APPLICANT: Bouquelere, Lydie
: TITLE OF INVENTION: Prostate cancer gene
: FILE REFERENCE: GENSET 018CP1
: CURRENT APPLICATION NUMBER: US/09/218,207
: CURRENT FILING DATE: 1998-12-22
: EARLIER APPLICATION NUMBER: 08/996,306
: EARLIER FILING DATE: 1997-12-22
: EARLIER APPLICATION NUMBER: 60/099,658
: EARLIER FILING DATE: 1998-09-09
: NUMBER OF SEQ ID NOS: 578
: SOFTWARE: Patent.pm
: SEQ ID NO 4
: LENGTH: 353
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: HELIX
: LOCATION: 1..33
: OTHER INFORMATION: Rao and Argos identification method, potential helix
: FEATURE:
: NAME/KEY: HELIX
: LOCATION: 4..20
: OTHER INFORMATION: Klein, Kanehisa and Delisi identification method, potent
: FEATURE:
: NAME/KEY: HELIX
: LOCATION: 4..24
: OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
: FEATURE:
: NAME/KEY: MYRISTATE
: LOCATION: 12..16
: OTHER INFORMATION: Prosite match
: FEATURE:
: NAME/KEY: HELIX
: LOCATION: 50..70
: OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
: FEATURE:
: NAME/KEY: CARBOHYD
: LOCATION: 57..59
: OTHER INFORMATION: Prosite match
: FEATURE:
: NAME/KEY: HELIX
: LOCATION: 76..96
: OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
: FEATURE:
: NAME/KEY: PHOSPHORYLATION
: LOCATION: 78
: OTHER INFORMATION: potential Tyrosine kinase site, Prosite match
: FEATURE:
: NAME/KEY: PHOSPHORYLATION
: LOCATION: 84
: OTHER INFORMATION: potential caseine kinase II site, Prosite match
: FEATURE:
: NAME/KEY: SITE
: LOCATION: 94..115
: OTHER INFORMATION: potential leucine zipper site, Prosite match
: FEATURE:
: NAME/KEY: MYRISTATE
: LOCATION: 119..123
: OTHER INFORMATION: potential site, Prosite match
```



FEATURE: PHOSPHORYLATION  
NAME/KEY: 133  
LOCATION: 133  
OTHER INFORMATION: potential protein kinase C, Prosite match  
FEATURE: PHOSPHORYLATION  
NAME/KEY: 147  
LOCATION: 147  
OTHER INFORMATION: potential caseine kinase II site, Prosite match  
FEATURE: PHOSPHORYLATION  
NAME/KEY: 194  
LOCATION: 194  
OTHER INFORMATION: potential protein kinase C, Prosite match  
FEATURE: PHOSPHORYLATION  
NAME/KEY: 215  
LOCATION: 215  
OTHER INFORMATION: potential Tyrosine kinase site, Prosite match  
FEATURE: SULFATATION  
NAME/KEY: 221  
LOCATION: 221  
OTHER INFORMATION: Prosite match  
FEATURE: PHOSPHORYLATION  
NAME/KEY: 223  
LOCATION: 223  
OTHER INFORMATION: potential cAMP and cGMP dependant protein kinase site, Prosite match  
FEATURE: PHOSPHORYLATION  
NAME/KEY: 235  
LOCATION: 235  
OTHER INFORMATION: potential caseine kinase II site, Prosite match  
FEATURE: PHOSPHORYLATION  
NAME/KEY: 306  
LOCATION: 306  
OTHER INFORMATION: potential protein kinase C, Prosite match  
FEATURE: HELIX  
NAME/KEY: 310..330  
LOCATION: 310..330  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F  
FEATURE: MYRISTATE  
NAME/KEY: 319..323  
LOCATION: 319..323  
OTHER INFORMATION: Prosite match  
FEATURE: MYRISTATE  
NAME/KEY: 323..327  
LOCATION: 323..327  
OTHER INFORMATION: Prosite match  
FEATURE: AMIDATION  
NAME/KEY: 329  
LOCATION: 329  
OTHER INFORMATION: Prosite match  
FEATURE: HELIX  
NAME/KEY: 333..353  
LOCATION: 333..353  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F  
FEATURE: MYRISTATE  
NAME/KEY: 341..345  
LOCATION: 341..345  
OTHER INFORMATION: Prosite match  
FEATURE: PHOSPHORYLATION  
NAME/KEY: 350  
LOCATION: 350  
OTHER INFORMATION: potential protein kinase C, Prosite match  
US-09-218-207-4

Query Match 100.0%; Score 228; DB 4; Length 353;  
Best local Similarity 100.0%; Pred. No. 1.6e-221;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLPSVLLGTAPTYVLAAGVWRLLSAFLPARFYQALDDRLCYVYQSWVLEFFENYTG 60  
|||||  
DB 1 MRYLPSVLLGTAPTYVLAAGVWRLLSAFLPARFYQALDDRLCYVYQSWVLEFFENYTG 60  
|||||  
QY 61 VQILLGDLPRKKNENIYLANHOSVDWIVADILAIRQNALGHRVYVLEGLKWLPLYGC 120  
|||||

DB 61 VQILLGDLPRKKNENIYLANHOSVDWIVADILAIRQNALGHRVYVLEGLKWLPLYGC 120  
QY 121 YFAOHGIIYKRSKAKNEKEMRNKLOSYDAGTPMYLVIFPEGTRNPEOTKVLASQAF 180  
|||||  
DB 121 YFAOHGIIYKRSKAKNEKEMRNKLOSYDAGTPMYLVIFPEGTRNPEOTKVLASQAF 180  
|||||  
QY 181 AARGILAVLKHVLTPIKATHVAFDCMKNYLDIYDVTVYVEGKDGG 228  
|||||  
DB 181 AARGILAVLKHVLTPIKATHVAFDCMKNYLDIYDVTVYVEGKDGG 228  
|||||

RESULT 6  
US-08-996-306-5  
; Sequence 5, Application US/08996306  
; Patent No. 5945522  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Chumakov, Ilya  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Bouguenoret, Lydie  
; TITLE OF INVENTION: Prostate cancer gene  
; NUMBER OF SEQUENCES: 68  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson & Bear  
; STREET: 501 West Broadway  
; City: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92101-3505  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Win95  
; SOFTWARE: Word  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/996,306  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israelien, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: GENSET. 018A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 364 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: SINGLE  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: PROTEIN  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: potential protein  
; LOCATION: 1..364  
; US-08-996-306-5

Query Match 100.0%; Score 228; DB 2; Length 364;  
Best local Similarity 100.0%; Pred. No. 1.7e-221;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLPSVLLGTAPTYVLAAGVWRLLSAFLPARFYQALDDRLCYVYQSWVLEFFENYTG 60  
|||||  
DB 12 MRYLPSVLLGTAPTYVLAAGVWRLLSAFLPARFYQALDDRLCYVYQSWVLEFFENYTG 71  
|||||  
QY 61 VQILLGDLPRKKNENIYLANHOSVDWIVADILAIRQNALGHRVYVLEGLKWLPLYGC 120  
|||||  
DB 72 VQILLGDLPRKKNENIYLANHOSVDWIVADILAIRQNALGHRVYVLEGLKWLPLYGC 131  
|||||  
QY 121 YFAOHGIIYKRSKAKNEKEMRNKLOSYDAGTPMYLVIFPEGTRNPEOTKVLASQAF 180  
|||||



```
Db      1 MRLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYCYQSMTLFFENYTG 60
QY      61 VOILLYGDLPRKKNENIYLANHOSYVDWIVADILAIRONALGHVRYVLEKGLKMLPLYGC 120
Db      61 VOILLYGDLPRKKNENIYLANHOSYVDWIVADILAIRONALGHVRYVLEKGLKMLPLYGC 120
QY      121 YFAOHGIIYKRSKAFNEKEMRNKLOSVDAGTPTMLVTFPEGSTRYNPQOTKVLASQAF 180
Db      121 YFAOHGIIYKRSKAFNEKEMRNKLOSVDAGTPTMLVTFPEGSTRYNPQOTKVLASQAF 180
QY      181 AAORG 185
Db      181 AAORG 185
```

```
RESULT 10
US-09-218-207-136
; Sequence 136, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bouqueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; FILE REFERENCE: GENSET.018CPI
; CURRENT APPLICATION NUMBER: US/09/218,207
; CURRENT FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 136
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 81..83
; OTHER INFORMATION: Box I
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 160..165
; OTHER INFORMATION: Box II
US-09-218-207-136
```

```
Query Match      81.1%; Score 185; DB 4; Length 185;
Best Local Similarity 100.0%; Pred. No. 1.6e-178;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 MRLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYCYQSMTLFFENYTG 60
Db      1 MRLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYCYQSMTLFFENYTG 60
QY      61 VOILLYGDLPRKKNENIYLANHOSYVDWIVADILAIRONALGHVRYVLEKGLKMLPLYGC 120
Db      61 VOILLYGDLPRKKNENIYLANHOSYVDWIVADILAIRONALGHVRYVLEKGLKMLPLYGC 120
QY      121 YFAOHGIIYKRSKAFNEKEMRNKLOSVDAGTPTMLVTFPEGSTRYNPQOTKVLASQAF 180
Db      121 YFAOHGIIYKRSKAFNEKEMRNKLOSVDAGTPTMLVTFPEGSTRYNPQOTKVLASQAF 180
QY      181 AAORG 185
Db      181 AAORG 185
```

```
RESULT 11
US-09-338-907-134
; Sequence 134, Application US/09338907
```

```
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bouqueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CPICP
; CURRENT APPLICATION NUMBER: US/09/338,907
; CURRENT FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218,207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 134
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 81..83
; OTHER INFORMATION: Box I
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 160..165
; OTHER INFORMATION: Box II
US-09-338-907-134
```

```
Query Match      81.1%; Score 185; DB 4; Length 315;
Best Local Similarity 100.0%; Pred. No. 2.7e-178;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 MRLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYCYQSMTLFFENYTG 60
Db      1 MRLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYCYQSMTLFFENYTG 60
QY      61 VOILLYGDLPRKKNENIYLANHOSYVDWIVADILAIRONALGHVRYVLEKGLKMLPLYGC 120
Db      61 VOILLYGDLPRKKNENIYLANHOSYVDWIVADILAIRONALGHVRYVLEKGLKMLPLYGC 120
QY      121 YFAOHGIIYKRSKAFNEKEMRNKLOSVDAGTPTMLVTFPEGSTRYNPQOTKVLASQAF 180
Db      121 YFAOHGIIYKRSKAFNEKEMRNKLOSVDAGTPTMLVTFPEGSTRYNPQOTKVLASQAF 180
QY      181 AAORG 185
Db      181 AAORG 185
```

```
RESULT 12
US-09-218-207-134
; Sequence 134, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bouqueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; FILE REFERENCE: GENSET.018CPI
; CURRENT APPLICATION NUMBER: US/09/218,207
; CURRENT FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
```

```
; SEQ ID NO 134
; LENGTH: 315
; TYPE: prt
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 81..83
; OTHER INFORMATION: Box I
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 160..165
; OTHER INFORMATION: Box II
US-09-218-207-134
```

```
Query Match      81.1%; Score 185; DB 4; Length 315;
Best Local Similarity 100.0%; Pred. No. 2.7e-178;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MRYLPSVVLGTAPTYVLAMGVWRLLSAFLPARFYQALDDRLCYQYSWVLEFFENYTG 60
   |||||||
Db 1 MRYLPSVVLGTAPTYVLAMGVWRLLSAFLPARFYQALDDRLCYQYSWVLEFFENYTG 60
QY 61 VOILLYGDLPRKKNENITYLANHOSTVDWIVADILAIROMNLGHVRYVLEKGLWMLPLYGC 120
   |||||||
Db 61 VOILLYGDLPRKKNENITYLANHOSTVDWIVADILAIROMNLGHVRYVLEKGLWMLPLYGC 120
QY 121 YFAOHGIIYKRSKAKNEKEMRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLSSQAF 180
   |||||||
Db 121 YFAOHGIIYKRSKAKNEKEMRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLSSQAF 180
QY 181 AAQRG 185
   |||||
Db 181 AAQRG 185
```

```
RESULT 13
US-09-338-907-135
; Sequence 135, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marla
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CPLCP
; CURRENT APPLICATION NUMBER: US/09/338,907
; EARLIER FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218,207
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 135
; LENGTH: 300
; TYPE: prt
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 81..83
; OTHER INFORMATION: Box I
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 160..165
; OTHER INFORMATION: Box II
US-09-338-907-135
```

Query Match 80.7%; Score 184; DB 4; Length 300;

Best Local Similarity 100.0%; Pred. No. 2.6e-177;  
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MRYLPSVVLGTAPTYVLAMGVWRLLSAFLPARFYQALDDRLCYQYSWVLEFFENYTG 60
   |||||||
Db 1 MRYLPSVVLGTAPTYVLAMGVWRLLSAFLPARFYQALDDRLCYQYSWVLEFFENYTG 60
QY 61 VOILLYGDLPRKKNENITYLANHOSTVDWIVADILAIROMNLGHVRYVLEKGLWMLPLYGC 120
   |||||||
Db 61 VOILLYGDLPRKKNENITYLANHOSTVDWIVADILAIROMNLGHVRYVLEKGLWMLPLYGC 120
QY 121 YFAOHGIIYKRSKAKNEKEMRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLSSQAF 180
   |||||||
Db 121 YFAOHGIIYKRSKAKNEKEMRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLSSQAF 180
QY 181 AAQR 184
   |||||
Db 181 AAQR 184

RESULT 14
US-09-218-207-135
; Sequence 135, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marla
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; FILE REFERENCE: GENSET.018CPL
; CURRENT APPLICATION NUMBER: US/09/218,207
; EARLIER FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 135
; LENGTH: 300
; TYPE: prt
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 81..83
; OTHER INFORMATION: Box I
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 160..165
; OTHER INFORMATION: Box II
US-09-218-207-135
```

Query Match 80.7%; Score 184; DB 4; Length 300;  
Best Local Similarity 100.0%; Pred. No. 2.6e-177;  
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MRYLPSVVLGTAPTYVLAMGVWRLLSAFLPARFYQALDDRLCYQYSWVLEFFENYTG 60
   |||||||
Db 1 MRYLPSVVLGTAPTYVLAMGVWRLLSAFLPARFYQALDDRLCYQYSWVLEFFENYTG 60
QY 61 VOILLYGDLPRKKNENITYLANHOSTVDWIVADILAIROMNLGHVRYVLEKGLWMLPLYGC 120
   |||||||
Db 61 VOILLYGDLPRKKNENITYLANHOSTVDWIVADILAIROMNLGHVRYVLEKGLWMLPLYGC 120
QY 121 YFAOHGIIYKRSKAKNEKEMRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLSSQAF 180
   |||||||
Db 121 YFAOHGIIYKRSKAKNEKEMRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLSSQAF 180
QY 181 AAQR 184
   |||||
Db 181 AAQR 184
```

RESULT 15  
US-09-338-907-133  
; Sequence 133, Application US/09338907  
; Patent No. 6265546  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Ilya, Chumakov  
; APPLICANT: Bouquelarel, Lydie  
; TITLE OF INVENTION: PROSTATE CANCER GENE  
; FILE REFERENCE: GENSET.18C1CP  
; CURRENT APPLICATION NUMBER: US/09/338, 907  
; CURRENT FILING DATE: 1999-06-23  
; EARLIER APPLICATION NUMBER: 08/996, 306  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: 60/099, 658  
; EARLIER FILING DATE: 1998-09-09  
; EARLIER APPLICATION NUMBER: 09/218, 207  
; EARLIER FILING DATE: 1998-12-22  
; NUMBER OF SEQ ID NOS: 578  
; SOFTWARE: Patent.pm  
; SEQ ID NO 133  
; LENGTH: 182  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: 81..83  
; OTHER INFORMATION: Box I  
US-09-338-907-133

Query Match 67.58; Score 154; DB 4; Length 182;  
Best Local Similarity 100.0%; Pred. No. 2.5e-147;  
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRYLPSVVLGTAPTYVLAMGVMRLSLAFLPARFYQALDDRLCYQOSMVLFFENYTG 60  
Db 1 MRYLPSVVLGTAPTYVLAMGVMRLSLAFLPARFYQALDDRLCYQOSMVLFFENYTG 60  
QY 61 VOILLYGDLPRKNENIYLANHOSVDMIVADILAIRONALGHVRYVLKEGLKWLPLYGC 120  
Db 61 VOILLYGDLPRKNENIYLANHOSVDMIVADILAIRONALGHVRYVLKEGLKWLPLYGC 120  
QY 121 YFAOHGIIYVKRSKAFNEKEMRNKLQSYVDAGTP 154  
Db 121 YFAOHGIIYVKRSKAFNEKEMRNKLQSYVDAGTP 154

Search completed: August 28, 2002, 11:19:22  
Job time: 305 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:25:19 ; Search time 305.3 Seconds  
(without alignments)  
262.861 Million cell updates/sec

Title: US-09-853-526-70

Perfect score: 228

Sequence: 1 MRYLPSVVLGTAPTYVLA.....NYLDATYVTVEYEGKDDGG 228

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 3502263 seqs, 351980561 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Pending\_Patents\_AA\_Main:\*

1: /cgn2\_6/ptodata/2/paa/PCITUS.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/paa/US06.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/paa/US07.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/paa/US08.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/paa/US081.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/paa/US082.COMB.pep:\*  
7: /cgn2\_6/ptodata/2/paa/US083.COMB.pep:\*  
8: /cgn2\_6/ptodata/2/paa/US084.COMB.pep:\*  
9: /cgn2\_6/ptodata/2/paa/US085.COMB.pep:\*  
10: /cgn2\_6/ptodata/2/paa/US086.COMB.pep:\*  
11: /cgn2\_6/ptodata/2/paa/US087.COMB.pep:\*  
12: /cgn2\_6/ptodata/2/paa/US088.COMB.pep:\*  
13: /cgn2\_6/ptodata/2/paa/US089.COMB.pep:\*  
14: /cgn2\_6/ptodata/2/paa/US090.COMB.pep:\*  
15: /cgn2\_6/ptodata/2/paa/US091.COMB.pep:\*  
16: /cgn2\_6/ptodata/2/paa/US092.COMB.pep:\*  
17: /cgn2\_6/ptodata/2/paa/US093.COMB.pep:\*  
18: /cgn2\_6/ptodata/2/paa/US094.COMB.pep:\*  
19: /cgn2\_6/ptodata/2/paa/US095.COMB.pep:\*  
20: /cgn2\_6/ptodata/2/paa/US096.COMB.pep:\*  
21: /cgn2\_6/ptodata/2/paa/US097.COMB.pep:\*  
22: /cgn2\_6/ptodata/2/paa/US098.COMB.pep:\*  
23: /cgn2\_6/ptodata/2/paa/US099.COMB.pep:\*  
24: /cgn2\_6/ptodata/2/paa/US100.COMB.pep:\*  
25: /cgn2\_6/ptodata/2/paa/US101.COMB.pep:\*  
26: /cgn2\_6/ptodata/2/paa/US60.COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	228	100.0	228	22	US-09-853-526-70
2	228	100.0	228	23	US-09-901-484A-70
3	228	100.0	228	26	US-60-099-658-70
4	228	100.0	353	1	PCT-US01-01431-59
5	228	100.0	353	1	PCT-US01-01431-79
6	228	100.0	353	1	PCT-US01-11988-1474
7	228	100.0	353	1	PCT-US01-11988-1475

8	228	100.0	353	18	US-09-488-725A-2736	Sequence 2736, Ap
9	228	100.0	353	22	US-09-833-245-1474	Sequence 1474, Ap
10	228	100.0	353	22	US-09-833-245-1475	Sequence 1475, Ap
11	228	100.0	353	22	US-09-853-526-4	Sequence 4, Appl1
12	228	100.0	353	23	US-09-901-484A-4	Sequence 4, Appl1
13	228	100.0	353	23	US-09-915-582-59	Sequence 59, Appl1
14	228	100.0	353	23	US-09-915-582-79	Sequence 79, Appl1
15	228	100.0	353	26	US-60-099-658-4	Sequence 4, Appl1
16	228	100.0	364	18	US-09-436-919-1	Sequence 1, Appl1
17	228	100.0	364	22	US-09-817-910-7	Sequence 7, Appl1
18	228	100.0	364	22	US-09-853-526-5	Sequence 5, Appl1
19	228	100.0	364	23	US-09-901-484A-5	Sequence 5, Appl1
20	228	100.0	364	26	US-60-099-658-5	Sequence 5, Appl1
21	185	81.1	185	22	US-09-853-526-136	Sequence 136, App
22	185	81.1	185	23	US-09-901-484A-136	Sequence 136, App
23	185	81.1	315	22	US-09-853-526-134	Sequence 134, App
24	185	81.1	315	23	US-09-901-484A-134	Sequence 134, App
25	184	80.7	300	22	US-09-853-526-135	Sequence 135, App
26	184	80.7	300	23	US-09-901-484A-135	Sequence 135, App
27	154	67.5	182	22	US-09-853-526-133	Sequence 133, App
28	154	67.5	182	23	US-09-901-484A-133	Sequence 133, App
29	144	63.2	269	1	PCT-US01-01327-47	Sequence 47, Appl
30	143	62.7	450	26	US-60-245-222-110	Sequence 110, App
31	142	62.3	1032	26	US-60-212-413-309	Sequence 309, App
32	142	62.3	1032	26	US-60-229-518-248	Sequence 248, App
33	127	55.7	372	18	US-09-488-725A-6308	Sequence 6308, App
34	105	46.1	291	22	US-09-853-526-127	Sequence 127, App
35	105	46.1	291	23	US-09-901-484A-127	Sequence 127, App
36	91	39.9	176	18	US-09-436-919-5	Sequence 5, Appl1
37	85	37.3	97	22	US-09-853-526-132	Sequence 132, App
38	85	37.3	97	23	US-09-901-484A-132	Sequence 132, App
39	81	35.5	92	17	US-09-316-123-124	Sequence 124, App
40	81	35.5	92	22	US-09-827-244-124	Sequence 124, App
41	74	32.5	261	22	US-09-853-526-128	Sequence 128, App
42	74	32.5	261	23	US-09-901-484A-128	Sequence 128, App
43	70	30.7	257	1	PCT-US01-08631-31719	Sequence 31719, A
44	70	30.7	980	1	PCT-US01-08631-42336	Sequence 42336, A
45	62	27.2	66	22	US-09-853-526-131	Sequence 131, App

#### ALIGNMENTS

RESULT 1

US-09-853-526-70

Sequence 70, Application US/09853526

GENERAL INFORMATION:

APPLICANT: Cohen, Daniel

APPLICANT: Blumenfeld, Marta

APPLICANT: Ilya, Chumakov

APPLICANT: Bougueleret, Lydie

TITLE OF INVENTION: PROSTATE CANCER GENE

FILE REFERENCE: GENSET.18CPICP

CURRENT APPLICATION NUMBER: US/09/853,526

CURRENT FILING DATE: 2001-05-11

PRIOR APPLICATION NUMBER: 09/238,907

PRIOR FILING DATE: 1999-06-23

PRIOR APPLICATION NUMBER: 08/996,306

PRIOR FILING DATE: 1997-12-22

PRIOR APPLICATION NUMBER: 60/099,658

PRIOR FILING DATE: 1998-09-09

PRIOR APPLICATION NUMBER: 09/218,207

PRIOR FILING DATE: 1998-12-22

NUMBER OF SEQ ID NOS: 578

SOFTWARE: Patent.pm

SEQ ID NO 70

LENGTH: 228

TYPE: PRT

ORGANISM: Homo sapiens

US-09-853-526-70

Query Match 100.0%, Score 228; DB 22; Length 228;

Best Local Similarity 100.0%; Pred. No. 6.3e-221;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MRYLLPSVVLGTAPTYVLMGWRLSAPFLPARFYQALDDRLXCVYQSMVLEFFENYTG 60
    |||||||
DB 1 MRYLLPSVVLGTAPTYVLMGWRLSAPFLPARFYQALDDRLXCVYQSMVLEFFENYTG 60
QY 61 VOILLYGDLPKKNENITYLANHSTVDWIVADILAIRQNLGHRVYLKGLKWLPLYGC 120
    |||||||
DB 61 VOILLYGDLPKKNENITYLANHSTVDWIVADILAIRQNLGHRVYLKGLKWLPLYGC 120
QY 121 YFAOHGIIYVRSRKAKEKEMRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASQAF 180
    |||||||
DB 121 YFAOHGIIYVRSRKAKEKEMRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASQAF 180
QY 181 AAORGLAVLKHVLTPIKATHVAFDCMKNYLDALYDVTYVEGKDDG 228
    |||||||
DB 181 AAORGLAVLKHVLTPIKATHVAFDCMKNYLDALYDVTYVEGKDDG 228
```

## RESULT 2

```
US-09-901-484A-70
; Sequence 70, Application US/09901484A
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate Cancer Gene
; FILE REFERENCE: GEN-T11XC3D2
; CURRENT APPLICATION NUMBER: US/09/901,484A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 08/996,306
; PRIOR FILING DATE: 1997-12-22
; PRIOR APPLICATION NUMBER: US 60/099,658
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: US 09/218,207
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: US 09/338,907
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: US 09/853,526
; PRIOR FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 70
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (5208)..(5208)
; OTHER INFORMATION: n = a, c, g, or t.
US-09-901-484A-70
```

Query Match 100.0%; Score 228; DB 23; Length 228;  
Best Local Similarity 100.0%; Pred. No. 6.3e-221;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MRYLLPSVVLGTAPTYVLMGWRLSAPFLPARFYQALDDRLXCVYQSMVLEFFENYTG 60
    |||||||
DB 1 MRYLLPSVVLGTAPTYVLMGWRLSAPFLPARFYQALDDRLXCVYQSMVLEFFENYTG 60
QY 61 VOILLYGDLPKKNENITYLANHSTVDWIVADILAIRQNLGHRVYLKGLKWLPLYGC 120
    |||||||
DB 61 VOILLYGDLPKKNENITYLANHSTVDWIVADILAIRQNLGHRVYLKGLKWLPLYGC 120
QY 121 YFAOHGIIYVRSRKAKEKEMRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASQAF 180
    |||||||
DB 121 YFAOHGIIYVRSRKAKEKEMRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASQAF 180
QY 181 AAORGLAVLKHVLTPIKATHVAFDCMKNYLDALYDVTYVEGKDDG 228
    |||||||
```

DB 181 AAORGLAVLKHVLTPIKATHVAFDCMKNYLDALYDVTYVEGKDDG 228

## RESULT 3

```
US-60-099-658-70
; Sequence 70, Application US/60099658
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Chumakov, Ilya
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; NUMBER OF SEQUENCES: 99
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/099,658
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: GENSET.018APR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 70:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: Protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
US-60-099-658-70
```

Query Match 100.0%; Score 228; DB 26; Length 228;  
Best Local Similarity 100.0%; Pred. No. 6.3e-221;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MRYLLPSVVLGTAPTYVLMGWRLSAPFLPARFYQALDDRLXCVYQSMVLEFFENYTG 60
    |||||||
DB 1 MRYLLPSVVLGTAPTYVLMGWRLSAPFLPARFYQALDDRLXCVYQSMVLEFFENYTG 60
QY 61 VOILLYGDLPKKNENITYLANHSTVDWIVADILAIRQNLGHRVYLKGLKWLPLYGC 120
    |||||||
DB 61 VOILLYGDLPKKNENITYLANHSTVDWIVADILAIRQNLGHRVYLKGLKWLPLYGC 120
QY 121 YFAOHGIIYVRSRKAKEKEMRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASQAF 180
    |||||||
DB 121 YFAOHGIIYVRSRKAKEKEMRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASQAF 180
QY 181 AAORGLAVLKHVLTPIKATHVAFDCMKNYLDALYDVTYVEGKDDG 228
    |||||||
DB 181 AAORGLAVLKHVLTPIKATHVAFDCMKNYLDALYDVTYVEGKDDG 228
```

## RESULT 4

```
PCT-US01-01431-59
; Sequence 59, Application PC/TUS0101431
; GENERAL INFORMATION:
```



```
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: 17 human secreted proteins
; FILE REFERENCE: PS723PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01431
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/231,968
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-01431-59
```

```
Query Match          100.0%; Score 228; DB 1; Length 353;
Best Local Similarity 100.0%; Pred. No. 9.3e-221;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MRYLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYCYQSQWVLEFFENYTG 60
   |||||||
Db 1 MRYLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYCYQSQWVLEFFENYTG 60
QY 61 VOILLYGDLPKKNENIYYLANHOSYDWMIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120
   |||||||
Db 61 VOILLYGDLPKKNENIYYLANHOSYDWMIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120
QY 121 YFAOHGGIYVKSAREKEMRNKLOSVDAGTIPMTLVIFPEGTRYNPEOTKVLASQAF 180
   |||||||
Db 121 YFAOHGGIYVKSAREKEMRNKLOSVDAGTIPMTLVIFPEGTRYNPEOTKVLASQAF 180
QY 181 AAORGLAVLKHVLTPIKATHVAFDCKMKNYLDIYDVTVVYEESKDDGG 228
   |||||||
Db 181 AAORGLAVLKHVLTPIKATHVAFDCKMKNYLDIYDVTVVYEESKDDGG 228
```

```
RESULT 5
PCT-US01-01431-79
; Sequence 79, Application PC/TUS0101431
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: 17 human secreted proteins
; FILE REFERENCE: PS723PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01431
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/231,968
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 79
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-01431-79
```

```
Query Match          100.0%; Score 228; DB 1; Length 353;
Best Local Similarity 100.0%; Pred. No. 9.3e-221;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MRYLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYCYQSQWVLEFFENYTG 60
   |||||||
Db 1 MRYLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYCYQSQWVLEFFENYTG 60
```

```
QY 61 VOILLYGDLPKKNENIYYLANHOSYDWMIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120
   |||||||
Db 61 VOILLYGDLPKKNENIYYLANHOSYDWMIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120
QY 121 YFAOHGGIYVKSAREKEMRNKLOSVDAGTIPMTLVIFPEGTRYNPEOTKVLASQAF 180
   |||||||
Db 121 YFAOHGGIYVKSAREKEMRNKLOSVDAGTIPMTLVIFPEGTRYNPEOTKVLASQAF 180
QY 181 AAORGLAVLKHVLTPIKATHVAFDCKMKNYLDIYDVTVVYEESKDDGG 228
   |||||||
Db 181 AAORGLAVLKHVLTPIKATHVAFDCKMKNYLDIYDVTVVYEESKDDGG 228
```

```
RESULT 6
PCT-US01-11988-1474
; Sequence 1474, Application PC/TUS0111988
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: PCT/US01/11988
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1474
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-11988-1474
```

```
Query Match          100.0%; Score 228; DB 1; Length 353;
Best Local Similarity 100.0%; Pred. No. 9.3e-221;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MRYLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYCYQSQWVLEFFENYTG 60
   |||||||
Db 1 MRYLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYCYQSQWVLEFFENYTG 60
QY 61 VOILLYGDLPKKNENIYYLANHOSYDWMIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120
   |||||||
Db 61 VOILLYGDLPKKNENIYYLANHOSYDWMIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120
QY 121 YFAOHGGIYVKSAREKEMRNKLOSVDAGTIPMTLVIFPEGTRYNPEOTKVLASQAF 180
   |||||||
Db 121 YFAOHGGIYVKSAREKEMRNKLOSVDAGTIPMTLVIFPEGTRYNPEOTKVLASQAF 180
QY 181 AAORGLAVLKHVLTPIKATHVAFDCKMKNYLDIYDVTVVYEESKDDGG 228
   |||||||
Db 181 AAORGLAVLKHVLTPIKATHVAFDCKMKNYLDIYDVTVVYEESKDDGG 228
```

```
RESULT 7
PCT-US01-11988-1475
; Sequence 1475, Application PC/TUS0111988
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: PCT/US01/11988
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
```

NUMBER OF SEQ ID NOS: 2267  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1475  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
PCT-US01-11988-1475

Query Match 100.0%; Score 228; DB 1; Length 353;  
Best Local Similarity 100.0%; Pred. No. 9.3e-221;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTAPTYVLAMGVWRLLSAFLPARFYQALDRLCYQSVWLFEEFENYTG 60  
DB 1 MRYLLPSVLLGTAPTYVLAMGVWRLLSAFLPARFYQALDRLCYQSVWLFEEFENYTG 60  
QY 61 VOILLYDGLPKKNENIYLLANHOSTVDMIVADILAIROMNLGHVRYVLEKGLKPLPGC 120  
DB 61 VOILLYDGLPKKNENIYLLANHOSTVDMIVADILAIROMNLGHVRYVLEKGLKPLPGC 120  
QY 121 YFAOHGIIYVRSKAKFENEKERNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASQAF 180  
DB 121 YFAOHGIIYVRSKAKFENEKERNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASQAF 180  
QY 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDALYDVTYVEGKDDG 228  
DB 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDALYDVTYVEGKDDG 228

RESULT 8  
US-09-488-725A-2736  
Sequence 2736, Application US/09488725A  
GENERAL INFORMATION:  
APPLICANT: Hyseq Inc  
TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides  
FILE REFERENCE: 784FLPCT  
CURRENT APPLICATION NUMBER: US/09/488,725A  
CURRENT FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: US/09/488,725  
PRIOR FILING DATE: 2000-01-21  
PRIOR APPLICATION NUMBER: US09/552,317  
PRIOR FILING DATE: 2000-04-25  
PRIOR APPLICATION NUMBER: US09/598,042  
PRIOR FILING DATE: 2000-06-20  
PRIOR APPLICATION NUMBER: US09/620,312  
PRIOR FILING DATE: 2000-07-19  
PRIOR APPLICATION NUMBER: US09/653,450  
PRIOR FILING DATE: 2000-08-31  
PRIOR APPLICATION NUMBER: US09/662,191  
PRIOR FILING DATE: 2000-09-14  
PRIOR APPLICATION NUMBER: US09/693,036  
PRIOR FILING DATE: 2000-10-19  
PRIOR APPLICATION NUMBER: US09/727,344  
PRIOR FILING DATE: 2000-11-29  
NUMBER OF SEQ ID NOS: 7144  
SOFTWARE: PL\_FL\_genes\_b Versions 1.0  
SEQ ID NO 2736  
LENGTH: 354  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-488-725A-2736

Query Match 100.0%; Score 228; DB 18; Length 353;  
Best Local Similarity 100.0%; Pred. No. 9.3e-221;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTAPTYVLAMGVWRLLSAFLPARFYQALDRLCYQSVWLFEEFENYTG 60  
DB 1 MRYLLPSVLLGTAPTYVLAMGVWRLLSAFLPARFYQALDRLCYQSVWLFEEFENYTG 60  
QY 61 VOILLYDGLPKKNENIYLLANHOSTVDMIVADILAIROMNLGHVRYVLEKGLKPLPGC 120

DB 61 VOILLYDGLPKKNENIYLLANHOSTVDMIVADILAIROMNLGHVRYVLEKGLKPLPGC 120  
QY 121 YFAOHGIIYVRSKAKFENEKERNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASQAF 180  
DB 121 YFAOHGIIYVRSKAKFENEKERNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASQAF 180  
QY 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDALYDVTYVEGKDDG 228  
DB 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDALYDVTYVEGKDDG 228

RESULT 9  
US-09-833-245-1474  
Sequence 1474, Application US/09833245  
GENERAL INFORMATION:  
APPLICANT: Human Genome Sciences, Inc.  
TITLE OF INVENTION: Albumin Fusion Proteins  
FILE REFERENCE: PFS46PCT  
CURRENT APPLICATION NUMBER: US/09/833,245  
CURRENT FILING DATE: 2001-04-12  
PRIOR APPLICATION NUMBER: 60/229,358  
PRIOR FILING DATE: 2000-04-12  
PRIOR APPLICATION NUMBER: 60/256,931  
PRIOR FILING DATE: 2000-12-21  
PRIOR APPLICATION NUMBER: 60/199,384  
PRIOR FILING DATE: 2000-04-25  
NUMBER OF SEQ ID NOS: 2267  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1474  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-833-245-1474

Query Match 100.0%; Score 228; DB 22; Length 353;  
Best Local Similarity 100.0%; Pred. No. 9.3e-221;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTAPTYVLAMGVWRLLSAFLPARFYQALDRLCYQSVWLFEEFENYTG 60  
DB 1 MRYLLPSVLLGTAPTYVLAMGVWRLLSAFLPARFYQALDRLCYQSVWLFEEFENYTG 60  
QY 61 VOILLYDGLPKKNENIYLLANHOSTVDMIVADILAIROMNLGHVRYVLEKGLKPLPGC 120  
DB 61 VOILLYDGLPKKNENIYLLANHOSTVDMIVADILAIROMNLGHVRYVLEKGLKPLPGC 120  
QY 121 YFAOHGIIYVRSKAKFENEKERNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASQAF 180  
DB 121 YFAOHGIIYVRSKAKFENEKERNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASQAF 180  
QY 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDALYDVTYVEGKDDG 228  
DB 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDALYDVTYVEGKDDG 228

RESULT 10  
US-09-833-245-1475  
Sequence 1475, Application US/09833245  
GENERAL INFORMATION:  
APPLICANT: Human Genome Sciences, Inc.  
TITLE OF INVENTION: Albumin Fusion Proteins  
FILE REFERENCE: PFS46PCT  
CURRENT APPLICATION NUMBER: US/09/833,245  
CURRENT FILING DATE: 2001-04-12  
PRIOR APPLICATION NUMBER: 60/229,358  
PRIOR FILING DATE: 2000-04-12  
PRIOR APPLICATION NUMBER: 60/256,931  
PRIOR FILING DATE: 2000-12-21  
PRIOR APPLICATION NUMBER: 60/199,384  
PRIOR FILING DATE: 2000-04-25  
NUMBER OF SEQ ID NOS: 2267

SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1475  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-853-245-1475

Query Match 100.0% Score 228: DB 22: Length 353;  
Best Local Similarity 100.0% Pred. No. 9.3e-221;  
Matches 228: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVVLGTAPTYVVLAMGVWRLLSAFLPARYQALDDRLRYCYQSMVLFEEFNNTG 60  
DB 1 MRYLLPSVVLGTAPTYVVLAMGVWRLLSAFLPARYQALDDRLRYCYQSMVLFEEFNNTG 60  
QY 61 VOILLXGDLPRKNENIYLANHOSYVDWIVADILAIRQNALGHVRYVLEKGLKMLPLYGC 120  
DB 61 VOILLXGDLPRKNENIYLANHOSYVDWIVADILAIRQNALGHVRYVLEKGLKMLPLYGC 120  
QY 121 YFAOHGIGIVKRSKAFNEKEMKNKLOSIVDAGTPMYLVLPFGSTRYNBPOTVLSASQAF 180  
DB 121 YFAOHGIGIVKRSKAFNEKEMKNKLOSIVDAGTPMYLVLPFGSTRYNBPOTVLSASQAF 180  
QY 181 AAORGLAVLKHVLPRIKATHVAFCMKNYLDAIVDYVYVYEGSKDDGG 228  
DB 181 AAORGLAVLKHVLPRIKATHVAFCMKNYLDAIVDYVYVYEGSKDDGG 228

RESULT 11  
US-09-853-526-4  
Sequence 4, Application US/09853526  
GENERAL INFORMATION:

APPLICANT: Cohen, Daniel  
APPLICANT: Blumenfeld, Marta  
APPLICANT: Iliya, Chumakov  
APPLICANT: Bougueleret, Lydie  
TITLE OF INVENTION: PROSTATE CANCER GENE  
FILE REFERENCE: GENSET.18CPLCP  
CURRENT APPLICATION NUMBER: US/09/853,526  
CURRENT FILING DATE: 2001-05-11  
PRIOR APPLICATION NUMBER: 09/338,907  
PRIOR FILING DATE: 1999-06-23  
PRIOR APPLICATION NUMBER: 08/996,306  
PRIOR FILING DATE: 1997-12-22  
PRIOR APPLICATION NUMBER: 60/099,658  
PRIOR FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: 09/218,207  
PRIOR FILING DATE: 1998-12-22  
NUMBER OF SEQ ID NOS: 578  
SOFTWARE: Patent.pm  
SEQ ID NO 4

LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 1..33  
OTHER INFORMATION: Rao and Argos identification method, potential helix  
NAME/KEY: HELIX  
LOCATION: 4..20  
OTHER INFORMATION: Klein, Kanehisa and Delisi identification method, potential  
NAME/KEY: HELIX  
LOCATION: 4..24  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F  
NAME/KEY: MYRISTATE  
LOCATION: 12..16  
OTHER INFORMATION: helix  
NAME/KEY: MYRISTATE  
LOCATION: 12..16  
OTHER INFORMATION: Prositte match  
NAME/KEY: HELIX  
LOCATION: 50..70  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F  
OTHER INFORMATION: helix

NAME/KEY: CARBOHYD  
LOCATION: 57..59  
OTHER INFORMATION: Prositte match  
NAME/KEY: HELIX  
LOCATION: 76..96  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 78  
OTHER INFORMATION: potential Tyrosine kinase site, Prositte match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 84  
OTHER INFORMATION: potential caseine kinase II site, Prositte match  
NAME/KEY: SITE  
LOCATION: 94..115  
OTHER INFORMATION: potential Leucine zipper site, Prositte match  
NAME/KEY: MYRISTATE  
LOCATION: 119..123  
OTHER INFORMATION: potential site, Prositte match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 133  
OTHER INFORMATION: potential protein kinase C, Prositte match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 147  
OTHER INFORMATION: potential caseine kinase II site, Prositte match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 194  
OTHER INFORMATION: potential protein kinase C, Prositte match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 215  
OTHER INFORMATION: potential Tyrosine kinase site, Prositte match  
NAME/KEY: SULFATATION  
LOCATION: 221  
OTHER INFORMATION: Prositte match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 233  
OTHER INFORMATION: potential cAMP and cGMP dependant protein kinase site,  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 235  
OTHER INFORMATION: potential caseine kinase II site, Prositte match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 306  
OTHER INFORMATION: potential protein kinase C, Prositte match  
NAME/KEY: HELIX  
LOCATION: 310..330  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
NAME/KEY: MYRISTATE  
LOCATION: 319..323  
OTHER INFORMATION: Prositte match  
NAME/KEY: MYRISTATE  
LOCATION: 323..327  
OTHER INFORMATION: Prositte match  
NAME/KEY: AMIDATION  
LOCATION: 329  
OTHER INFORMATION: Prositte match  
NAME/KEY: HELIX  
LOCATION: 333..353  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
NAME/KEY: MYRISTATE  
LOCATION: 341..345  
OTHER INFORMATION: Prositte match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 350  
OTHER INFORMATION: potential protein kinase C, Prositte match  
US-09-853-526-4

Query Match 100.0% Score 228: DB 22: Length 353;  
Best Local Similarity 100.0% Pred. No. 9.3e-221;  
Matches 228: Conservative 0; Mismatches 0; Indels 0; Gaps 0;





NAME/KEY: potential Transmembrane helix  
LOCATION: 1..33  
IDENTIFICATION METHOD: Rao and Argos method  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 4..20  
IDENTIFICATION METHOD: Klein, Kanehisa and Delisi method  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 4..24  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komaromy  
IDENTIFICATION METHOD: and Wall method  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 12  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 50..70  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komaromy and Wall method  
FEATURE:  
NAME/KEY: potential N-glycosylation site  
LOCATION: 57  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 76..96  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komaromy and Wall method  
FEATURE:  
NAME/KEY: potential Tyrosine kinase phosphorylation site  
LOCATION: 78  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Casein kinase II phosphorylation site  
LOCATION: 84  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Leucine zipper pattern  
LOCATION: 94..115  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 119  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 133  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Casein kinase II phosphorylation site  
LOCATION: 147  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 194  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Tyrosine kinase phosphorylation site  
LOCATION: 215  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Tyrosine sulfatation site  
LOCATION: 221  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential CAMP- and cAMP-dependent protein kinase phosphorylation site  
LOCATION: 233  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Casein kinase II phosphorylation site  
LOCATION: 235  
IDENTIFICATION METHOD: prosite match  
FEATURE:

NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 306  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 310..330  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komaromy and Wall method  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 319  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 323  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Amladation site  
LOCATION: 329  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 333..353  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komaromy and Wall method  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 341  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 350  
IDENTIFICATION METHOD: prosite match  
US-60-099-658-4

Query Match 100.0%; Score 228; DB 26; Length 353;  
Best Local Similarity 100.0%; Pred No. 9.3e-221;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRYLLPSVVLGTAPTYVLAMGWRLSAFLPARFYQALDDRLCYQSMVLEFFENYTG 60  
Db 1 MRYLLPSVVLGTAPTYVLAMGWRLSAFLPARFYQALDDRLCYQSMVLEFFENYTG 60  
QY 61 VOILLYGDLEPRKKNENIYLANHOSYVDIADILAIQNALGHVRYLAKGLKWLPLYGC 120  
Db 61 VOILLYGDLEPRKKNENIYLANHOSYVDIADILAIQNALGHVRYLAKGLKWLPLYGC 120  
QY 121 YFAOHGSIYVRSKAKFNEKEKRNKLOSYVDAGTPMYLVIPPEGTRVNPEDOTKVLASQAF 180  
Db 121 YFAOHGSIYVRSKAKFNEKEKRNKLOSYVDAGTPMYLVIPPEGTRVNPEDOTKVLASQAF 180  
QY 181 AAOGLAVLKHVLPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDG 228  
Db 181 AAOGLAVLKHVLPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDG 228

Search completed: August 28, 2002, 11:25:20  
Job time: 498 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:26:50 ; Search time 83.88 Seconds

(without alignments)  
657.933 Million cell updates/sec

Title: US-09-853-526-70

Perfect score: 228  
Sequence: 1 MRYLPSVVLGTAFTYVLA.....NYLDAIVDVVEGKDDGS 228

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 747981 seqs, 242050750 residues

Word size : 0

Total number of hits satisfying chosen parameters: 747981

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Pending\_Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	228	100.0	364	US-10-184-648-63	Sequence 63, Appl
2	144	63.2	269	US-10-074-045-47	Sequence 47, Appl
3	144	63.2	353	US-09-629-469A-13028	Sequence 13028, A
4	9	3.9	135	US-10-184-648-52	Sequence 52, Appl
5	9	3.9	157	US-10-184-648-66	Sequence 66, Appl
6	9	3.9	240	US-09-895-913A-184	Sequence 184, App
7	8	3.5	236	US-10-053-853A-1538	Sequence 1538, Ap
8	8	3.5	262	US-09-791-537-5820	Sequence 5820, Ap
9	8	3.5	459	US-09-791-537-110883	Sequence 110883,
10	8	3.5	483	US-09-791-537-110642	Sequence 110642,
11	7	3.1	44	US-09-826-734A-142	Sequence 142, Appl
12	7	3.1	77	US-09-548-936C-17	Sequence 17, Appl
13	7	3.1	96	US-09-791-537-150979	Sequence 150979,
14	7	3.1	117	US-09-791-537-31675	Sequence 31675, A
15	7	3.1	126	US-09-791-537-63528	Sequence 63528, A
16	7	3.1	194	US-09-791-537-53250	Sequence 53250, A
17	7	3.1	195	US-10-164-966-9	Sequence 9, Appl
18	7	3.1	195	US-10-184-648-51	Sequence 51, Appl
19	7	3.1	195	US-10-184-648-67	Sequence 67, Appl
20	7	3.1	195	US-10-184-648-86	Sequence 86, Appl
21	7	3.1	208	PCT-US02-10824-132	Sequence 132, App
22	7	3.1	208	US-09-791-537-57367	Sequence 57367, A
23	7	3.1	208	US-09-791-537-99121	Sequence 99121, A
24	7	3.1	208	US-09-791-537-128069	Sequence 128069,
25	7	3.1	208	US-09-791-537-142901	Sequence 142901,
26	7	3.1	208	US-09-791-537-143774	Sequence 143774,

#### ALIGNMENTS

27	7	3.1	208	6	US-10-096-327-2	Sequence 2, Appl1
28	7	3.1	208	6	US-10-179-131-7124	Sequence 7124, Ap
29	7	3.1	208	6	US-10-189-360-5	Sequence 5, Appl1
30	7	3.1	208	6	US-10-189-360-7	Sequence 7, Appl1
31	7	3.1	208	6	US-10-189-360-8	Sequence 8, Appl1
32	7	3.1	208	6	US-10-138-158-18	Sequence 18, Appl
33	7	3.1	224	5	US-09-791-537-148811	Sequence 148811,
34	7	3.1	255	6	US-10-155-881-22428	Sequence 22428, A
35	7	3.1	262	5	US-09-791-537-75154	Sequence 751, App
36	7	3.1	263	6	US-10-155-881-18654	Sequence 18654, A
37	7	3.1	269	6	US-10-053-853A-1244	Sequence 1244, Ap
38	7	3.1	270	6	US-10-211-364-1026	Sequence 1026, Ap
39	7	3.1	286	7	US-60-360-039-1733	Sequence 1733, Ap
40	7	3.1	364	5	US-09-791-537-42339	Sequence 42339, A
41	7	3.1	400	5	US-09-791-537-17833	Sequence 17833, A
42	7	3.1	409	5	US-09-791-537-116428	Sequence 116428,
43	7	3.1	441	5	US-09-791-537-94112	Sequence 94112, A
44	7	3.1	445	5	US-09-791-537-117991	Sequence 117991,
45	7	3.1	499	5	US-09-791-537-107567	Sequence 107567,

RESULT 1  
US-10-184-648-63  
Sequence 63, Application US/10184648  
GENERAL INFORMATION:  
APPLICANT: Meyers, Rachel E.  
APPLICANT: Williamson, Mark  
APPLICANT: Tsai, Fong-Ying  
APPLICANT: Hunter, John J.  
APPLICANT: Macbeth, Kyle J.  
APPLICANT: Rudolph-Owen, Laura A.  
APPLICANT: Leiby, Kevin R.  
APPLICANT: Kapeller-Libermann, Rosana  
APPLICANT: Olandt, Peter J.  
TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE FAMILY MEMBERS AND USES THEREOF  
FILE REFERENCE: 10448-192001  
CURRENT APPLICATION NUMBER: US/10/184, 648  
CURRENT FILING DATE: 2002-06-27  
PRIOR APPLICATION NUMBER: US 09/815, 028  
PRIOR FILING DATE: 2001-03-22  
PRIOR APPLICATION NUMBER: PCT/US01/09358  
PRIOR FILING DATE: 2001-03-22  
PRIOR APPLICATION NUMBER: US 60/191, 964  
PRIOR FILING DATE: 2000-03-24  
PRIOR APPLICATION NUMBER: US 09/801, 220  
PRIOR FILING DATE: 2001-03-07  
PRIOR APPLICATION NUMBER: PCT/US01/07269  
PRIOR FILING DATE: 2001-03-07  
PRIOR APPLICATION NUMBER: US 60/187, 456  
PRIOR FILING DATE: 2000-03-07  
PRIOR APPLICATION NUMBER: US 09/816, 714  
PRIOR FILING DATE: 2001-03-23  
PRIOR APPLICATION NUMBER: PCT/US01/09468  
PRIOR FILING DATE: 2001-03-23  
PRIOR APPLICATION NUMBER: US 60/191, 865  
PRIOR FILING DATE: 2000-03-24  
PRIOR APPLICATION NUMBER: US 09/844, 948  
PRIOR FILING DATE: 2001-04-27  
PRIOR APPLICATION NUMBER: PCT/US01/13805  
PRIOR FILING DATE: 2001-04-27  
PRIOR APPLICATION NUMBER: US 60/200, 604  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: US 09/861, 164  
PRIOR FILING DATE: 2001-05-18  
PRIOR APPLICATION NUMBER: PCT/US01/16292  
PRIOR FILING DATE: 2001-05-18  
PRIOR APPLICATION NUMBER: US 60/205, 408  
PRIOR FILING DATE: 2000-05-19  
PRIOR APPLICATION NUMBER: US 09/883, 060  
PRIOR FILING DATE: 2001-06-15

```

; PRIOR APPLICATION NUMBER: PCT/US01/19138
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,079
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/962,678
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: PCT/US01/29963
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/235,044
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 09/973,457
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: US 60/238,849
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 10/072,285
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: PCT/US02/03736
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/267,494
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: US 09/817,910
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: PCT/US01/09633
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/192,092
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: US 09/842,528
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: PCT/US01/40607
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 60/199,500
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/882,836
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19543
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/211,730
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/882,872
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19153
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,077
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-184-648-63
```

```

Query Match      100.0%; Score 228; DB 6; Length 364;
Best Local Similarity 100.0%; Pred. No. 2.9e-225;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 MRYLLPSVLLGTAPTYVLAAGVWRLSAPFLPARFYQALDDRLCYQOSWLPFFENTG 60
    |||||
DB 12 MRYLLPSVLLGTAPTYVLAAGVWRLSAPFLPARFYQALDDRLCYQOSWLPFFENTG 71
    |||||
QY 61 VOILLYDGLPRKNENIYLANHCSIVDMIVADILAIROMALGHVRYVLEGLKWLPLYGC 120
    |||||
DB 72 VOILLYDGLPRKNENIYLANHCSIVDMIVADILAIROMALGHVRYVLEGLKWLPLYGC 131
    |||||
QY 121 YFAHGIGIYKRSAKFEKEKRNKLOSYVDAGTPTMYLVIPEEGTRYNPEQTKVLSASQAF 180
    |||||
DB 132 YFAHGIGIYKRSAKFEKEKRNKLOSYVDAGTPTMYLVIPEEGTRYNPEQTKVLSASQAF 191
    |||||
QY 181 AAGRGVLVLAHVLPRIKATHVAFDCKMKNYLDATYDVTVYVEGKDDG 228
    |||||
DB 192 AAGRGVLVLAHVLPRIKATHVAFDCKMKNYLDATYDVTVYVEGKDDG 239
    |||||
```

```

RESULT 2
US-10-074-045-47
; Sequence 47; Application US/10074045
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT221C1
; CURRENT APPLICATION NUMBER: US/10/074,045
; CURRENT FILING DATE: 2002-02-14
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 47
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-074-045-47
```

```

Query Match      63.2%; Score 144; DB 6; Length 269;
Best Local Similarity 100.0%; Pred. No. 3.1e-139;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 85 TVDMIVADILAIROMALGHVRYVLEGLKWLPLYGCTFAHGIGIYKRSAKFEKRNK 144
    |||||
DB 1 TVDMIVADILAIROMALGHVRYVLEGLKWLPLYGCTFAHGIGIYKRSAKFEKRNK 60
    |||||
QY 145 LOSYVDAGTPTMYLVIPEEGTRYNPEQTKVLSASQAFNAOGLAVLKHVLPRIKATHVAF 204
    |||||
DB 61 LOSYVDAGTPTMYLVIPEEGTRYNPEQTKVLSASQAFNAOGLAVLKHVLPRIKATHVAF 120
    |||||
QY 205 DCKMKNYLDATYDVTVYVEGKDDG 228
    |||||
DB 121 DCKMKNYLDATYDVTVYVEGKDDG 144
    |||||
```

```

RESULT 3
US-09-629-469A-13028
; Sequence 13028; Application US/09629469A
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: SAITO, KAORI
; APPLICANT: YAMAMOTO, JUNICHI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: WAKAMATSU, AI
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: OTSUKI, TETSUJI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE
; FILE REFERENCE: 084335/0123
; CURRENT APPLICATION NUMBER: US/09/629,469A
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: JP 1999-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 1999-300253
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183,322
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 19025
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13028
```



LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-629-469A-13028

Query Match 63.2%; Score 144; DB 5; Length 353;  
Best Local Similarity 100.0%; Pred. No. 4e-139;  
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MRYLLPSVVLGTAPTYVLAWGVWRLLSAFLPAREYQALDDRLCYVQSMVLFEEFNNTG 60  
Db 1 MRYLLPSVVLGTAPTYVLAWGVWRLLSAFLPAREYQALDDRLCYVQSMVLFEEFNNTG 60  
Oy 61 VOILLYGDLPRKKENITVIANHOSYVDVADILAIROMALGHVRYLKEGKMLPLXGC 120  
Db 61 VOILLYGDLPRKKENITVIANHOSYVDVADILAIROMALGHVRYLKEGKMLPLXGC 120  
Oy 121 YFAOHGIYVKRSKAKFNEKEMRNK 144  
Db 121 YFAOHGIYVKRSKAKFNEKEMRNK 144

RESULT 4  
US-10-184-648-52  
Sequence 52, Application US/10184648

GENERAL INFORMATION:  
APPLICANT: Meyers, Rachel E.  
APPLICANT: Williamson, Mark  
APPLICANT: Tsal, Fong-Ying  
APPLICANT: Hunter, John J.  
APPLICANT: Macbeth, Kyle J.  
APPLICANT: Rudolph-Owen, Laura A.  
APPLICANT: Leiby, Kevin R.  
APPLICANT: Kapeller-Liebermann, Rosana  
APPLICANT: Olandt, Peter J.  
TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE FAMILY MEMBERS AND USES THEREOF  
FILE REFERENCE: 10448-192001  
CURRENT APPLICATION NUMBER: US/10/184,648  
CURRENT FILING DATE: 2002-06-27  
PRIOR APPLICATION NUMBER: US 09/815,028  
PRIOR FILING DATE: 2001-03-22  
PRIOR APPLICATION NUMBER: PCT/US01/09358  
PRIOR FILING DATE: 2001-03-22  
PRIOR APPLICATION NUMBER: US 60/191,964  
PRIOR FILING DATE: 2000-03-24  
PRIOR APPLICATION NUMBER: US 09/801,220  
PRIOR FILING DATE: 2001-03-07  
PRIOR APPLICATION NUMBER: PCT/US01/07269  
PRIOR FILING DATE: 2001-03-07  
PRIOR APPLICATION NUMBER: US 60/187,456  
PRIOR FILING DATE: 2000-03-07  
PRIOR APPLICATION NUMBER: US 09/816,714  
PRIOR FILING DATE: 2001-03-23  
PRIOR APPLICATION NUMBER: PCT/US01/09468  
PRIOR FILING DATE: 2001-03-23  
PRIOR APPLICATION NUMBER: US 60/191,865  
PRIOR FILING DATE: 2000-03-24  
PRIOR APPLICATION NUMBER: US 09/844,948  
PRIOR FILING DATE: 2001-04-27  
PRIOR APPLICATION NUMBER: PCT/US01/13805  
PRIOR FILING DATE: 2001-04-27  
PRIOR APPLICATION NUMBER: US 60/200,604  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: US 09/861,164  
PRIOR FILING DATE: 2001-05-18  
PRIOR APPLICATION NUMBER: PCT/US01/16292  
PRIOR FILING DATE: 2001-05-18  
PRIOR APPLICATION NUMBER: US 60/205,408  
PRIOR FILING DATE: 2000-05-19  
PRIOR APPLICATION NUMBER: US 09/883,060  
PRIOR FILING DATE: 2001-06-15  
PRIOR APPLICATION NUMBER: PCT/US01/19138

PRIOR FILING DATE: 2001-06-15  
PRIOR APPLICATION NUMBER: US 60/212,079  
PRIOR FILING DATE: 2000-06-15  
PRIOR APPLICATION NUMBER: US 09/962,678  
PRIOR FILING DATE: 2001-09-25  
PRIOR APPLICATION NUMBER: PCT/US01/29963  
PRIOR FILING DATE: 2001-09-25  
PRIOR APPLICATION NUMBER: US 60/235,044  
PRIOR FILING DATE: 2000-09-25  
PRIOR APPLICATION NUMBER: US 09/973,457  
PRIOR FILING DATE: 2001-10-09  
PRIOR APPLICATION NUMBER: US 60/238,849  
PRIOR FILING DATE: 2000-10-06  
PRIOR APPLICATION NUMBER: US 10/072,285  
PRIOR FILING DATE: 2002-02-08  
PRIOR APPLICATION NUMBER: PCT/US02/03736  
PRIOR FILING DATE: 2002-02-08  
PRIOR APPLICATION NUMBER: US 60/267,494  
PRIOR FILING DATE: 2001-02-08  
PRIOR APPLICATION NUMBER: US 09/817,910  
PRIOR FILING DATE: 2001-03-26  
PRIOR APPLICATION NUMBER: PCT/US01/09633  
PRIOR FILING DATE: 2001-03-26  
PRIOR APPLICATION NUMBER: US 60/192,092  
PRIOR FILING DATE: 2000-03-24  
PRIOR APPLICATION NUMBER: US 09/842,528  
PRIOR FILING DATE: 2001-04-25  
PRIOR APPLICATION NUMBER: PCT/US01/40607  
PRIOR FILING DATE: 2001-04-25  
PRIOR APPLICATION NUMBER: US 60/199,500  
PRIOR FILING DATE: 2000-04-25  
PRIOR APPLICATION NUMBER: US 09/882,836  
PRIOR FILING DATE: 2001-06-15  
PRIOR APPLICATION NUMBER: PCT/US01/19543  
PRIOR FILING DATE: 2001-06-15  
PRIOR APPLICATION NUMBER: US 60/211,730  
PRIOR FILING DATE: 2000-06-15  
PRIOR APPLICATION NUMBER: US 09/882,872  
PRIOR FILING DATE: 2001-06-15  
PRIOR APPLICATION NUMBER: PCT/US01/19153  
PRIOR FILING DATE: 2001-06-15  
PRIOR APPLICATION NUMBER: US 60/212,077  
PRIOR FILING DATE: 2000-06-15  
NUMBER OF SEQ ID NOS: 90  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 52  
LENGTH: 135  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: consensus sequence  
US-10-184-648-52

Query Match 3.9%; Score 9; DB 6; Length 135;  
Best Local Similarity 100.0%; Pred. No. 0.48;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 157 LVIFPEGTR 165  
Db 85 LVIFPEGTR 93

RESULT 5  
US-10-184-648-66  
Sequence 66, Application US/10184648  
GENERAL INFORMATION:  
APPLICANT: Meyers, Rachel E.  
APPLICANT: Williamson, Mark  
APPLICANT: Tsal, Fong-Ying  
APPLICANT: Hunter, John J.  
APPLICANT: Macbeth, Kyle J.  
APPLICANT: Rudolph-Owen, Laura A.

; APPLICANT: Leiby, Kevin R.  
 ; APPLICANT: Kapeller-Liebermann, Rosana  
 ; APPLICANT: Olandt, Peter J.  
 ; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE FAMILY MEMBERS AND USES THEREOF  
 ; FILE REFERENCE: 10448-192001  
 ; CURRENT APPLICATION NUMBER: US/10/184,648  
 ; CURRENT FILING DATE: 2002-06-27  
 ; PRIOR APPLICATION NUMBER: US 09/815,028  
 ; PRIOR FILING DATE: 2001-03-22  
 ; PRIOR APPLICATION NUMBER: PCT/US01/09358  
 ; PRIOR FILING DATE: 2001-03-22  
 ; PRIOR APPLICATION NUMBER: US 60/191,964  
 ; PRIOR FILING DATE: 2000-03-24  
 ; PRIOR APPLICATION NUMBER: US 09/801,220  
 ; PRIOR FILING DATE: 2001-03-07  
 ; PRIOR APPLICATION NUMBER: PCT/US01/07269  
 ; PRIOR FILING DATE: 2001-03-07  
 ; PRIOR APPLICATION NUMBER: US 60/187,456  
 ; PRIOR FILING DATE: 2000-03-07  
 ; PRIOR APPLICATION NUMBER: US 09/816,714  
 ; PRIOR FILING DATE: 2001-03-23  
 ; PRIOR APPLICATION NUMBER: PCT/US01/09468  
 ; PRIOR FILING DATE: 2001-03-23  
 ; PRIOR APPLICATION NUMBER: US 60/191,865  
 ; PRIOR FILING DATE: 2000-03-24  
 ; PRIOR APPLICATION NUMBER: US 09/844,948  
 ; PRIOR FILING DATE: 2001-04-27  
 ; PRIOR APPLICATION NUMBER: PCT/US01/13805  
 ; PRIOR FILING DATE: 2001-04-27  
 ; PRIOR APPLICATION NUMBER: US 60/200,604  
 ; PRIOR FILING DATE: 2000-04-28  
 ; PRIOR APPLICATION NUMBER: US 09/861,164  
 ; PRIOR FILING DATE: 2001-05-18  
 ; PRIOR APPLICATION NUMBER: PCT/US01/16292  
 ; PRIOR FILING DATE: 2001-05-18  
 ; PRIOR APPLICATION NUMBER: US 60/205,408  
 ; PRIOR FILING DATE: 2000-05-19  
 ; PRIOR APPLICATION NUMBER: US 09/883,060  
 ; PRIOR FILING DATE: 2001-06-15  
 ; PRIOR APPLICATION NUMBER: PCT/US01/19138  
 ; PRIOR FILING DATE: 2001-06-15  
 ; PRIOR APPLICATION NUMBER: US 60/212,079  
 ; PRIOR FILING DATE: 2000-06-15  
 ; PRIOR APPLICATION NUMBER: US 09/962,678  
 ; PRIOR FILING DATE: 2001-09-25  
 ; PRIOR APPLICATION NUMBER: PCT/US01/29963  
 ; PRIOR FILING DATE: 2001-09-25  
 ; PRIOR APPLICATION NUMBER: US 60/235,044  
 ; PRIOR FILING DATE: 2000-09-25  
 ; PRIOR APPLICATION NUMBER: US 09/973,457  
 ; PRIOR FILING DATE: 2001-10-09  
 ; PRIOR APPLICATION NUMBER: US 60/238,849  
 ; PRIOR FILING DATE: 2000-10-06  
 ; PRIOR APPLICATION NUMBER: US 10/072,285  
 ; PRIOR FILING DATE: 2002-02-08  
 ; PRIOR APPLICATION NUMBER: PCT/US02/03736  
 ; PRIOR FILING DATE: 2002-02-08  
 ; PRIOR APPLICATION NUMBER: US 60/267,494  
 ; PRIOR FILING DATE: 2001-02-08  
 ; PRIOR APPLICATION NUMBER: US 09/817,910  
 ; PRIOR FILING DATE: 2001-03-26  
 ; PRIOR APPLICATION NUMBER: PCT/US01/09633  
 ; PRIOR FILING DATE: 2001-03-26  
 ; PRIOR APPLICATION NUMBER: US 60/192,092  
 ; PRIOR FILING DATE: 2000-03-24  
 ; PRIOR APPLICATION NUMBER: US 09/842,528  
 ; PRIOR FILING DATE: 2001-04-25  
 ; PRIOR APPLICATION NUMBER: PCT/US01/40607  
 ; PRIOR FILING DATE: 2001-04-25  
 ; PRIOR APPLICATION NUMBER: US 60/199,500  
 ; PRIOR FILING DATE: 2000-04-25  
 ; PRIOR APPLICATION NUMBER: US 09/882,836  
 ; PRIOR FILING DATE: 2001-06-15

; PRIOR APPLICATION NUMBER: PCT/US01/19543  
 ; PRIOR FILING DATE: 2001-06-15  
 ; PRIOR APPLICATION NUMBER: US 60/211,730  
 ; PRIOR FILING DATE: 2000-06-15  
 ; PRIOR APPLICATION NUMBER: US 09/882,872  
 ; PRIOR FILING DATE: 2001-06-15  
 ; PRIOR APPLICATION NUMBER: PCT/US01/19153  
 ; PRIOR FILING DATE: 2001-06-15  
 ; PRIOR APPLICATION NUMBER: US 60/212,077  
 ; PRIOR FILING DATE: 2000-06-15  
 ; NUMBER OF SEQ ID NOS: 90  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 66  
 ; LENGTH: 157  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: consensus sequence  
 ; US-10-184-648-66

Query Match 3.9%; Score 9; DB 6; Length 157;  
 Best Local Similarity 100.0%; Pred. No. 0.55;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 211 LDAIYDVT 219  
 Db 4 LDAIYDVT 12

RESULT 6  
 US-09-895-913A-184  
 ; Sequence 184, Application US/09895913A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kleenothous, Harold  
 ; APPLICANT: Al-Garawi, Amal  
 ; APPLICANT: Miller, Charles  
 ; APPLICANT: Tomb, Jean Francois  
 ; APPLICANT: Oomen, Raymond P.  
 ; TITLE OF INVENTION: Identification of Polynucleotides  
 ; TITLE OF INVENTION: Encoding Novel Helicobacter Polypeptides in the Helicobacter  
 ; TITLE OF INVENTION: Genome  
 ; FILE REFERENCE: 06132/043002  
 ; CURRENT APPLICATION NUMBER: US/09/895,913A  
 ; CURRENT FILING DATE: 2001-06-29  
 ; PRIOR APPLICATION NUMBER: US 08/881,227  
 ; PRIOR FILING DATE: 1997-06-24  
 ; NUMBER OF SEQ ID NOS: 368  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 184  
 ; LENGTH: 240  
 ; TYPE: PRT  
 ; ORGANISM: Helicobacter pylori  
 ; US-09-895-913A-184

Query Match 3.9%; Score 9; DB 5; Length 240;  
 Best Local Similarity 100.0%; Pred. No. 0.82;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 LVTFPEGTR 165  
 Db 142 LVTFPEGTR 150

RESULT 7  
 US-10-053-853A-1538  
 ; Sequence 1538, Application US/10053853A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HAYASHI, Hideo  
 ; APPLICANT: SHINAGAWA, Hideo  
 ; APPLICANT: MAKINO, KOZO  
 ; APPLICANT: HAYASHI, Tetsuya

```

; APPLICANT: OHNISHI, Makoto
; APPLICANT: HATTORI, Masahira
; APPLICANT: KUROKAWA, Ken
; TITLE OF INVENTION: Polynucleotide molecules and polypeptides specific to Enterohemorrhagic E. coli O157:H7
; FILE REFERENCE: 2002-0060A/MC/01704
; CURRENT APPLICATION NUMBER: US/10/053,853A
; PRIOR FILING DATE: 2002-05-28
; PRIOR APPLICATION NUMBER: JP2001-112010
; NUMBER OF SEQ ID NOS: 1866
; SEQ ID NO 1538
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Escherichia coli O157:H7
US-10-053-853A-1538

Query Match
Best Local Similarity 3.5%; Score 8; DB 6; Length 236;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 LPPRIKAT 200
Db 227 LPPRIKAT 234

RESULT 8
US-09-791-537-5820
; Sequence 5820, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5820
; LENGTH: 262
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-09-791-537-5820

Query Match
Best Local Similarity 3.5%; Score 8; DB 5; Length 262;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 HYLTPRIK 198
Db 202 HYLTPRIK 209

RESULT 9
US-09-791-537-110883
; Sequence 110883, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 110883
; LENGTH: 459
```

```

; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-09-791-537-110883

Query Match
Best Local Similarity 3.5%; Score 8; DB 5; Length 459;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 167 NPEOTKYL 174
Db 292 NPEOTKYL 299

RESULT 10
US-09-791-537-110642
; Sequence 110642, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 110642
; LENGTH: 483
; TYPE: PRT
; ORGANISM: Pisum sativum
US-09-791-537-110642

Query Match
Best Local Similarity 3.5%; Score 8; DB 5; Length 483;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 AGTPMYLY 158
Db 184 AGTPMYLY 191

RESULT 11
US-09-826-734A-142
; Sequence 142, Application US/09826734A
; GENERAL INFORMATION:
; APPLICANT: Fernandes, Elma R
; APPLICANT: Mishra, Vishnu S
; APPLICANT: Leach, Martin D
; APPLICANT: Shimkets, Richard A
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Kekuda, Ramesh
; TITLE OF INVENTION: Novel Polynucleotides And Polypeptides Encoded Thereby
; FILE REFERENCE: 15966-754
; CURRENT APPLICATION NUMBER: US/09/826,734A
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/195,576
; PRIOR FILING DATE: 2000-04-26
; NUMBER OF SEQ ID NOS: 264
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 142
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-826-734A-142

Query Match
Best Local Similarity 3.1%; Score 7; DB 5; Length 44;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

OY 159 IPEPCTR 165  
|||||||  
Db 3 IPEPCTR 9

RESULT 12  
US-09-548-936C-17  
; Sequence 17, Application US/09548936C  
; GENERAL INFORMATION:  
; APPLICANT: Gopez, Leonel Jorge  
; APPLICANT: Sarras, Jan  
; APPLICANT: Claesson-Weish, Lena  
; APPLICANT: Heidlin, Carl-Jemrik  
; TITLE OF INVENTION: PTPL BINDING AGENTS  
; FILE REFERENCE: L0461/7084  
; CURRENT APPLICATION NUMBER: US/09/548,936C  
; CURRENT FILING DATE: 2000-04-13  
; PRIOR APPLICATION NUMBER: US 09/100,804  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: US 08/596,291  
; PRIOR FILING DATE: 1994-09-01  
; PRIOR APPLICATION NUMBER: US 08/115,573  
; PRIOR FILING DATE: 1993-09-01  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: PatentIn Version 3.0  
; SEQ ID NO 17  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-548-936C-17

Query Match 3.18; Score 7; DB 5; Length 77;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 125 HGGIYVK 131  
|||||||  
Db 20 HGGIYVK 26

RESULT 13  
US-09-791-537-150979  
; Sequence 150979, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Blomomix, Inc.  
; APPLICANT: Danzer, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn Version 3.0  
; SEQ ID NO 150979  
; LENGTH: 96  
; TYPE: PRT  
; ORGANISM: pdb 3PDZA  
; FEATURE:  
; NAME/KEY: misc-feature  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: X is an unknown amino acid  
US-09-791-537-150979

Query Match 3.18; Score 7; DB 5; Length 96;  
Best Local Similarity 100.0%; Pred. No. 39;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 125 HGGIYVK 131  
|||||||  
Db 32 HGGIYVK 38

RESULT 14  
US-09-791-537-31675  
; Sequence 31675, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Blomomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn Version 3.0  
; SEQ ID NO 31675  
; LENGTH: 117  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-791-537-31675

Query Match 3.18; Score 7; DB 5; Length 117;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 125 HGGIYVK 131  
|||||||  
Db 43 HGGIYVK 49

RESULT 15  
US-09-791-537-63528  
; Sequence 63528, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Blomomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn Version 3.0  
; SEQ ID NO 63528  
; LENGTH: 126  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-791-537-63528

Query Match 3.18; Score 7; DB 5; Length 126;  
Best Local Similarity 100.0%; Pred. No. 50;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 125 HGGIYVK 131  
|||||||  
Db 52 HGGIYVK 58

Search completed: August 28, 2002, 11:26:51  
Job time: 564 sec

8

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:27:18 ; Search time 21.78 Seconds  
(without alignments)  
647.104 Million cell updates/sec

Title: US-09-853-526-5

Perfect score: 364

Sequence: 1 MLSTLVHTFSMRYLPSV.....YVNTWIVGTLLGCLMTVRIKA 364

Scoring table:

OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size : 0

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	353	97.0	353	P1CE_HUMAN	Q9NQ22 homo sapien
2	32	8.8	354	P1CE_MOUSE	Q9D1E8 mus musculu
3	9	2.5	237	P1SC_HELPY	O9Z1N8 heliobacte
4	9	2.5	240	P1SC_HELPY	O25603 heliobacte
5	8	2.2	83	V187_BPT3	P10302 bacterioph
6	8	2.2	83	V187_BPT3	P03788 bacterioph
7	8	2.2	396	V187_BPT3	P10302 bacterioph
8	2.2	611	1	HSCA_BUCAL	P57660 buchnera ap
9	7	1.9	57	YUDO_ECOLI	P58638 escherichia
10	7	1.9	194	PURL_IACCA	P35853 lactobacill
11	7	1.9	208	HBCF_CERAE	O09118 cercoptithe
12	7	1.9	208	HBCF_HUMAN	O99075 homo sapien
13	7	1.9	208	HBCF_PIG	O01580 sus scrofa
14	7	1.9	208	HBCF_PIG	O06175 ratus norv
15	7	1.9	250	HXB4_MOUSE	P10284 mus musculu
16	7	1.9	250	HXB4_MOUSE	O59188 borrelia bu
17	7	1.9	250	HXB4_MOUSE	O23449 caenorhabdi
18	7	1.9	251	HXB4_HUMAN	P17483 homo sapien
19	7	1.9	286	PROC_YEAST	P32263 saccharomyc
20	7	1.9	308	PLSC_COCONU	Q24270 cocco nuctif
21	7	1.9	311	MPCP_YEAST	P23641 saccharomyc
22	7	1.9	423	ENO_METJA	O60173 methanococ
23	7	1.9	445	XVLA_BACSU	P04788 bacillus su
24	7	1.9	516	GUX1_PHACH	P13860 phanerocha
25	7	1.9	520	CMCH_NOCCLA	O51080 nocardia la
26	7	1.9	532	GSI_NEBCR	O51080 nocardia la
27	7	1.9	677	Y593_TREPA	O83602 treponema p
28	7	1.9	860	CH12_COCCIM	P54197 coccidioid
29	7	1.9	904	Y002_CAEEL	O09328 caenorhabdi
30	7	1.9	1018	ST31_MOUSE	O99NM1 mus musculu
31	7	1.9	1019	ST31_HUMAN	O99NM1 mus sapien
32	7	1.9	1026	BGAL_STRTF	P23989 streptococ
33	7	1.9	1043	SYL_METTH	O27428 methanobact

34	7	1.9	1044	1	SYL_METTH	P26499 methanobact
35	7	1.9	2485	1	PTND_HUMAN	O12923 homo sapien
36	6	1.6	25	1	PCN1_PACGO	P82423 pachycondyl
37	6	1.6	25	1	PCW2_PACGO	P82424 pachycondyl
38	6	1.6	71	1	YAAA_BACSU	P05650 bacillus su
39	6	1.6	76	1	B811_SCHCO	P87442 schizophyll
40	6	1.6	79	1	PSK3_ARATH	O9M2Y0 arabidopsis
41	6	1.6	83	1	PRRP_RAT	P81378 ratus norv
42	6	1.6	83	1	RL23_HALLA	O06842 halobacteri
43	6	1.6	87	1	YIMC_BPPH1	P10435 bacterioph
44	6	1.6	88	1	PMRD_ECOLI	P37590 escherichia
45	6	1.6	88	1	Y40L_RHISN	P55597 rhizobium s

#### ALIGNMENTS

RESULT 1  
P1CE\_HUMAN STANDARD: PRT; 353 AA.  
ID P1CE\_HUMAN  
AC Q9NU02; Q9B0G4;  
DT 01-MAR-2002 (Rel. 41, Created)  
DT 01-MAR-2002 (Rel. 41, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE 1-acyl-sn-glycerol-3-phosphate acyltransferase epsilon (EC 2.3.1.51)  
DE (1-AGP acyltransferase 5) (1-AGPAT 5) (Lysophosphatidic acid  
DE O-acyltransferase-epsilon) (LPAAT-epsilon) (1-acylglycerol-3-phosphate  
DE AGPAT5.  
GN Homo sapiens (Human).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RT Leung D.W.;  
RT "Cloning and expression of LPAAT-epsilon."  
RT Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Cohen D., Chumakov I., Blumenfeld M., Bouguetelaret L.;  
RN Patent number WO9932644, 01-JUL-1999.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Amnydata;  
RX MEDLINE=21154917; PubMed=11230166;  
RA Wiedmann S., Well B., Wellenreuther R., Gassenhuber J., Glassl S.,  
RA Ansoerger W., Boecker M., Bloeker H., Bauersachs S., Blum H.,  
RA Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,  
RA Mewes H.-W., Othenwaelder B., Obermaier B., Tampe J., Heubner D.,  
RA Wandt R., Korn B., Klein M., Poustka A.;  
RT "Towards a catalog of human genes and proteins: sequencing and  
RT analysis of 500 novel complete protein coding human cDNAs."  
RN Genome Res. 11:422-435(2001).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Placenta;  
RA Tsugei T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
RA Nishikawa T., Nagai K., Sugano S., Shiraori A., Sudo H.,  
RA Takatsuna M., Hosoliri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
RA Watanabe S., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
RA Yamamoto S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
RA Yamamoto J., Makatsus A., Nakamura Y., Nagahari K., Masuno Y.,  
RA Nimomiyu K., Iwayanagi T.;  
RT "NEBD human cDNA sequencing project."  
RP Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: CONVERTS LYSOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC  
CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION (BY  
CC SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate -  
CC CoA + 1,2-diacyl-sn-glycerol 3-phosphate.  
CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
 CC ACYLTRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF375789; AAK54809.1; ALT\_INIT.  
 DR EMBL: AL136587; CAB6552.1; ALT\_INIT.  
 DR EMBL: AK002072; BAB92069.1; -  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KW Phospholipid biosynthesis; Transferase; Acyltransferase;  
 KW Transmembrane.  
 FT TRANSMEM. 7 23 POTENTIAL.  
 FT TRANSMEM. 44 65 POTENTIAL.  
 FT TRANSMEM. 334 351 POTENTIAL.  
 FT CONFLICT 145 145 L -> V (IN REF. 2).  
 SQ SEQUENCE 353 AA; 40813 MW; A05B1FA246CE1B64 CRC64;

Query Match 97.0%; Score 353; DB 1; Length 353;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 MRYLLPSVLLGTAPTYLVANGVRLLSAFLPARFYQALDRLKCVQSMVLFRENTG 71  
 DB 1 MRYLLPSVLLGTAPTYLVANGVRLLSAFLPARFYQALDRLKCVQSMVLFRENTG 60  
 QY 72 VOILLYGDLPKKNKIYYLANHSTVDIVADILAIRONALGHRYVKEGKMLPLYGC 131  
 DB 61 VOILLYGDLPKKNKIYYLANHSTVDIVADILAIRONALGHRYVKEGKMLPLYGC 120  
 QY 132 YFAOHGGIYVRSKAFKFNKEMRNKLOSYVDAQTPMYLVIFPEGTRVNEQOTKVLASQAF 191  
 DB 121 YFAOHGGIYVRSKAFKFNKEMRNKLOSYVDAQTPMYLVIFPEGTRVNEQOTKVLASQAF 180  
 QY 192 AAOGLAVLKHVLTPIRIKATVAFDCKKNYDAITYDVVYEGDDGQRESEPTMEFL 251  
 DB 181 AAOGLAVLKHVLTPIRIKATVAFDCKKNYDAITYDVVYEGDDGQRESEPTMEFL 240  
 QY 252 CKESPKIHIDRIIDKIVPEEQEHMRRLHREFEIKDKMLIEFYESPDPERRRRPFCKS 311  
 DB 241 CKESPKIHIDRIIDKIVPEEQEHMRRLHREFEIKDKMLIEFYESPDPERRRRPFCKS 300  
 QY 312 VNSKLSIKTKLPSMLILSGLTAGMLMTDAGRKLVTWYVYGTLLGCLWVTKA 364  
 DB 301 VNSKLSIKTKLPSMLILSGLTAGMLMTDAGRKLVTWYVYGTLLGCLWVTKA 353

RESULT 2  
 PLSC\_MOUSE STANDARD: PRT; 354 AA.  
 AC 09DIE8;  
 DT 01-MAR-2002 (Rel. 41, Created)  
 DT 01-MAR-2002 (Rel. 41, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE 1-acyl-sn-glycerol-3-phosphate acyltransferase epsilon (EC 2.3.1.51)  
 DE 1-acyl acyltransferase 5) (1-AGPAT 5) (lysophosphatidic acid  
 DE acyltransferase-epsilon) (LPAAT-epsilon) (1-acylglycerol-3-phosphate  
 DE O-acyltransferase 5).  
 GN AGPAT5.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Embryo;

RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinaigawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Aikawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
 RA Aikawa K., Izawa M., Nishi K., Kiyosawa H., Rondo S., Yamana I.,  
 RA Saito T., Okazaki Y., Gotohori T., Bono H., Katsukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schirral L.M., Staehli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carlini P., de Bonaudo M.F.,  
 RA Brownstein M.J., Bolt C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Saeki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmink L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai J., Kohzuki S.,  
 RA Hayashizaki Y.,  
 RA "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 CC -1- FUNCTION: CONVERTS LYSOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC  
 CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION (BY  
 CC SIMILARITY)  
 CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate =  
 CC CoA + 1,2-diacyl-sn-glycerol 3-phosphate.  
 CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
 CC ACYLTRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AK003649; BAB22915.1; -  
 DR MGD: MGI:1915880; Agpat5.  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KW Phospholipid biosynthesis; Transferase; Acyltransferase;  
 KW Transmembrane.  
 FT TRANSMEM. 7 29 POTENTIAL.  
 FT TRANSMEM. 44 66 POTENTIAL.  
 FT TRANSMEM. 335 352 POTENTIAL.  
 SQ SEQUENCE 354 AA; 40943 MW; 447EEF924B91E800 CRC64;

Query Match 8.8%; Score 32; DB 1; Length 354;  
 Best Local Similarity 100.0%; Pred. No. 1; 6e-24;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 185 LSASQAFPAORGLAVLKHVLTPIRIKATVAFD 216  
 DB 174 LSASQAFPAORGLAVLKHVLTPIRIKATVAFD 205  
 RESULT 3  
 PLSC\_HELPJ STANDARD: PRT; 237 AA.  
 AC 09ZJNB;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE 1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51) (1-AGP  
 DE acyltransferase) (1-AGPAT) (lysophosphatidic acid acyltransferase)  
 DE (LPAAT).  
 GN PLSC OR JHP1267.  
 OS Helicobacter pylori J99 (Campylobacter pylori J99).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;



```
OC Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RA Alm R.A., Ling L.-S.L., Molt D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Camel G.,
RA Tummulo P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori."
RL Nature 397:176-180(1999).
CC -1- FUNCTION: CONVERTS LYSOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC
CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION.
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate =
CC COA + 1,2-diacyl-sn-glycerol 3-phosphate.
CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).
CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE
CC ACYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sid.ch).
CC -----
DR EMBL: AE001550; AAD06852.1; -
DR InterPro: IPR002123; Acyltransferase.
DR Pfam: PF01553; Acyltransferase: 1.
KW Phospholipid biosynthesis; Transferase; Acyltransferase;
KW Inner membrane; Complete proteome.
SQ SEQUENCE 237 AA; 27188 MW; E10F517D42A1731F CRC64;

Query Match 2.5%; Score 9; DB 1; Length 237;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 LVIFPESTR 176
ID |||||||
DB 142 LVIFPESTR 150

RESULT 4
PLSC_HELPY STANDARD; PRT; 240 AA.
ID PLSC_HELPY
OC 025903;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51) (1-AGP
DE acyltransferase) (1-AGPAT) (Lysophosphatidic acid acyltransferase)
DE (LIPAT).
GN PLSC OR HP1348.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=26695 / ATCC 700392;
RA MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klein H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khakh H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Uterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
```

```
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori."
RL Nature 388:539-547(1997).
CC -1- FUNCTION: CONVERTS LYSOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC
CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION.
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate =
CC COA + 1,2-diacyl-sn-glycerol 3-phosphate.
CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).
CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE
CC ACYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000636; AAD08393.1; -
DR TIGR: HP1348; -
DR InterPro: IPR002123; Acyltransferase.
DR Pfam: PF01553; Acyltransferase: 1.
KW Phospholipid biosynthesis; Transferase; Acyltransferase;
KW Inner membrane; Complete proteome.
SQ SEQUENCE 240 AA; 27745 MW; 22BD5D0EB190BBD CRC64;

Query Match 2.5%; Score 9; DB 1; Length 240;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 LVIFPESTR 176
ID |||||||
DB 142 LVIFPESTR 150

RESULT 5
V187_BP73 STANDARD; PRT; 83 AA.
ID V187_BP73
AC P10302;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-MAR-1989 (Rel. 10, Last annotation update)
DE Gene 18.7 protein.
GN 18.7.
OS Bacteriophage T3.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC T7-like phages.
OX NCBI_TaxID=10759;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=86209997; PubMed=3010556;
RA Yamada M., Fujisawa H., Kato H., Hamada K., Mingawa T.;
RT "Cloning and sequencing of the genetic right end of bacteriophage T3
RL DNA."
RL Virology 151:350-361(1986).
RN [2]
RP ERRATUM.
RA Yamada M., Fujisawa H., Kato H., Hamada K., Mingawa T.;
RL Virology 154:246-246(1986).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

DR EMBL: M14784; AAA92527.1; -.  
DR PIR: E23476; WBBPT3.  
SQ SEQUENCE 83 AA; 9393 MW; 63128984CBAF0531 CRC64;

Query Match  
Best Local Similarity 100.0%; Score 8; DB 1; Length 83;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 316 LSIKKTLP 323  
Db 22 LSIKKTLP 29

RESULT 6  
V187\_BPT7  
ID V187\_BPT7 STANDARD; PRT; 83 AA.  
AC P03786;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 01-MAR-1989 (Rel. 10, Last annotation update)  
DE Gene 18.7 protein.  
GN 18.7.  
OS Bacteriophage T7.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;  
OC T7-like phages.  
OX NCBI\_TaxID=10760;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=83241725; PubMed=6864790;  
RA Dunn J.J., Studier F.W.;  
RT "Complete nucleotide sequence of bacteriophage T7 DNA and the  
RT locations of T7 genetic elements."  
RL J. Mol. Biol. 166:477-535(1983).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC  
DR EMBL: V01146; CAA24439.1; -.  
DR PIR: A04413; WBBPG7.  
DR PIR: S42337; S42337.  
SQ SEQUENCE 83 AA; 9326 MW; B3DB055DE79A53C5 CRC64;

Query Match  
Best Local Similarity 100.0%; Score 8; DB 1; Length 83;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 316 LSIKKTLP 323  
Db 22 LSIKKTLP 29

RESULT 7  
YD18\_YEAST  
ID YD18\_YEAST STANDARD; PRT; 396 AA.  
AC Q12185;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Hypothetical 45.9 kDa protein in KCS1-GCV1 intergenic region.  
GN YOR018C OR Y09335.04C OR PZE596.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=49332;  
RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN=S288C / AB972;  
RA Dedman K., Brown D., Hamlyn N., Bowman S., Barrell B.G.,  
RA Rajandream M.A.;  
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.

RM [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97051598; PubMed=8896275;  
RA Elde L.G., Sander C., Prydz H.;  
RT "Sequencing and analysis of a 35.4 kb region on the right arm of  
RT chromosome IV from Saccharomyces cerevisiae reveal 23 open reading  
RT frames."  
RL Yeast 12:1085-1090(1996).  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
CC ACYLTRANSFERASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC  
DR EMBL: Z74314; CAA98838.1; -.  
DR EMBL: X95966; CAA65210.1; -.  
DR EMBL: Z49770; CAA89843.1; -.  
DR SGD: S0002425; YDR018C.  
DR InterPro: IPR002123; Acyltransferase.  
DR Pfam: PF01535; Acyltransferase; 1.  
KW Hypothetical protein; Phospholipid biosynthesis; Transferase;  
KW Acyltransferase; Transmembrane.  
FT TRANSMEM 27 47  
FT TRANSMEM 69 89 POTENTIAL.  
FT TRANSMEM 123 143 POTENTIAL.  
FT TRANSMEM 372 392 POTENTIAL.  
SQ SEQUENCE 396 AA; 45938 MW; 9F55AB56C8FD44DD CRC64;

Query Match  
Best Local Similarity 100.0%; Score 8; DB 1; Length 396;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 222 LDAIVDVT 229  
Db 265 LDAIVDVT 272

RESULT 8  
HSCA\_BUCAI  
ID HSCA\_BUCAI STANDARD; PRT; 611 AA.  
AC P57660;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Chaperone protein hscA homolog.  
GN HSCA OR BU605.  
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum  
OS symbiotic bacterium).  
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.  
OX NCBI\_TaxID=118099;  
RN [1]  
RP SEQUENCE FROM N.A.  
RP STRAIN=TOKYO 1998;  
RX MEDLINE=20445173; PubMed=10993077;  
RA Shigenobu S., Matanabe H., Hattori M., Sakaki Y., Ishikawa H.;  
RT "Genome sequence of the endocellular bacterial symbiont of aphids  
RT Buchnera sp. APS."  
RL Nature 407:81-86(2000).  
CC -1- FUNCTION: PROBABLE CHAPERONE. HAS A LOW INTRINSIC ATPASE ACTIVITY  
CC WHICH IS MARKEDLY STIMULATED BY HSCB (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AP001119; BAB13289.1; -  
 DR Interpro: IPR001023; HSP70.  
 DR Pfam: PF00012; HSP70.1.  
 DR PRINTS: PR00301; HEATSHOCK70.  
 DR PROSITE: PS00297; HSP70\_1; 1.  
 DR PROSITE: PS00329; HSP70\_2; 1.  
 DR PROSITE: PS01036; HSP70\_3; 1.  
 KM Chaperone; ATP-binding; Complete proteome.  
 SQ SEQUENCE 611 AA; 69837 MW; 32B567C53073082A CRC64;

Query Match 2.2%; Score 8; DB 1; Length 611;  
 Best Local Similarity 100.0%; Pred. No. 5.1;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 RYLLPSV 20  
 |||||  
 DB 42 RYLLPSV 49

## RESULT 9

ID YJDO\_ECOLI STANDARD; PRT; 57 AA.  
 AC P58038;

DT 16-OCT-2001 (Rel. 40, Last Created)

DT 16-OCT-2001 (Rel. 41, Last annotation update)

DE 01-MAR-2002 (Rel. 41, Last annotation update)

DE Hypothetical protein yj40.

GN YJDO OR B4128.1 OR Z5731 OR EC55110.1.

OS Escherichia coli, and

OS Escherichia coli O157:H7.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Escherichia.

OX NCBI\_Taxid=562, 83334;

RP SEQUENCE FROM N.A.  
 (1)

RC MEDLINE-97426617; PubMed-9278503;

RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,

RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,

RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,

RA Mau B., Shao Y.;

RT "The complete genome sequence of Escherichia coli K-12.";

RL Science 277:1453-1474(1997).

RN SEQUENCE FROM N.A.  
 (2)

RC STRAIN-0157:H7 / EDL933 / ATCC 700927;

RA MEDLINE-21074935; PubMed-11206551;

RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,

RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,

RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,

RA Grodock E.J., Davis N.W., Lim A., Dimalanta E.T., Potamowski K.,

RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,

RA Welch R.A., Blatner F.R.;

RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";

RL Nature 409:529-533(2001).

RN SEQUENCE FROM N.A.  
 (3)

RC STRAIN-0157:H7 / RIMD 0509952;

RA MEDLINE-21156231; PubMed-11258796;

RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,

RA Han C.-G., Ohsudo E., Nakayama K., Murata T., Tanaka M., Tode T.,

RA Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,

RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;

RT "Complete genome sequence of enterohemorrhagic Escherichia coli

RT O157:H7 and genomic comparison with a laboratory strain K-12.";  
 RL DNA Res. 8:11-22(2001).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC -1- SIMILARITY: STRONG, TO E.COLI YDCX.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: AE000485; -; NOT\_ANNOTATED\_CDS.

DR EMBL: AP002568; -; NOT\_ANNOTATED\_CDS.

DR EcoGene: EGI4342; yj40.

DE Hypothetical protein; Transmembrane; Complete proteome.

KW Hypothetical protein; Transmembrane; Complete proteome.

FT TRANSMEM 7 27 POTENTIAL.

FT TRANSMEM 37 57 POTENTIAL.

SQ SEQUENCE 57 AA; 6555 MW; A3670A19500F75D6 CRC64;

OY 36 RLISAF 42  
 |||||  
 DB 33 RLISAF 39

## RESULT 10

ID PURL\_LACCA STANDARD; PRT; 194 AA.  
 AC P35853;

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE Amidophosphoribosyltransferase precursor (EC 2.4.2.14) (Glutamine  
 DE phosphoribosylpyrophosphate amidotransferase) (ATASE) (GPATASE)

DE (Fragment).

DE PURP.

GN Lactobacillus casei.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;

OC Lactobacillus.

OX NCBI\_Taxid=1582;

RN SEQUENCE FROM N.A.  
 (1)

RX MEDLINE-93012962; PubMed-1398079;

RA Gu Z.-M., Martindale D.W., Lee B.H.;

RT "Isolation and complete sequence of the purL gene encoding PCAM

RT synthase II in Lactobacillus casei.";

RL Gene 119:123-126(1992).

RN SEQUENCE FROM N.A.  
 (2)

RP ERRATUM.

RX MEDLINE-94040790; PubMed-8224889;

RA Gu Z.-M., Martindale D.W., Lee B.H.;

RT "Isolation and complete sequence of the purL gene encoding PCAM

RT synthase II in Lactobacillus casei.";

RL Gene 133:147-147(1993).

CC -1- CATALYTIC ACTIVITY: 5-phospho-beta-D-ribose + diphosphate +

CC L-glutamate -> L-glutamine + 5-phospho-alpha-D-ribose 1-diphosphate

CC + H(2)O.

CC -1- PATHWAY: FIRST STEP IN DE NOVO PURINE BIOSYNTHESIS

CC -1- SIMILARITY: THE GATASE DOMAIN BELONGS TO TYPE-2 GLUTAMINE

CC AMIDOTRANSFERASES.

CC -1- SIMILARITY: IN THE C-TERMINAL SECTION, BELONGS TO THE

CC PURINE/PYRIMIDINE PHOSPHORIBOSYLTRANSFERASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL: M85265; AAC36948.1; .  
 DR PIR: PC1136; PC1136.  
 DR HSSP: P00497; 1A00.  
 DR MEROPS: C44.001; .  
 DR InterPro: IPR000583; GATase\_2.  
 DR InterPro: IPR002375; Pur\_Pyr\_pr\_transf.  
 DR Pfam: PF00310; GATase\_2; 1.  
 DR PROSITE: PS00103; PUR\_PYR\_PR\_TRANSFER; PARTIAL.  
 DR PROSITE: PS00443; GATASE\_TYPE\_II; 1.  
 DR Purine biosynthesis; Transferase; Glycosyltransferase.  
 KW PROPEP 1 1..  
 FT CHAIN 12 >194  
 FT ACT\_SITE 12 12  
 FT NON\_TER 194 194  
 FT SEQUENCE 194 AA; 21144 MW; 4A788BC5365D5EC CRC64;

Query Match 1.9%; Score 7; DB 1; Length 194;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 111 ALGHVRY 117  
 |||||  
 Db 79 ALGHVRY 85

RESULT 11  
 HBGF\_CERAE STANDARD; PRT; 208 AA.  
 ID HBGF\_CERAE 009118;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Heparin-binding EGF-like growth factor precursor (HB-EGF) (HBEGF)  
 DE (Diphtheria toxin receptor) (DT-R).  
 GN DTR OR HEGFL.  
 OS Cercopithecus aethiops (Green monkey) (Grivet).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Cercopithecus.  
 NC NCBL\_TaxID=9534;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92298386; PubMed=1606612;  
 RA Neglich J.G., Metherall J.E., Russel D.W., Eldels L.;  
 RT "Expression cloning of a diphtheria toxin receptor: Identify with a  
 heparin-binding EGF-like growth factor precursor.";  
 RL Cell 69:1051-1061(1992).  
 RN [2]  
 RP TOXIN-BINDING DOMAIN.  
 RX MEDLINE=95126975; PubMed=7826391;  
 RA Hooper K.P., Eldels L.;  
 RT "Localization of a critical diphtheria toxin-binding domain to the C-  
 terminus of the mature heparin-binding EGF-like growth factor region  
 of the diphtheria toxin receptor.";  
 RL Biochem. Biophys. Res. Commun. 206:710-717(1995).  
 CC -1- FUNCTION: MAY BE INVOLVED IN MACROPHAGE-MEDIATED CELLULAR  
 PROLIFERATION. IT IS MITOGENIC FOR FIBROBLASTS AND SMOOTH MUSCLE  
 BUT NOT ENDOTHELIAL CELLS. IT IS ABLE TO BIND EGF RECEPTORS WITH  
 HIGHER AFFINITY THAN EGF ITSELF AND IS A FAR MORE POTENT MITOGEN  
 FOR SMOOTH MUSCLE CELLS THAN EGF (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MATURE HB-EGF IS  
 RELEASED INTO THE EXTRACELLULAR SPACE AND PROBABLY BINDS TO A  
 RECEPTOR (BY SIMILARITY).  
 CC -1- PTM: O-GLYCOSYLATED (BY SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

DR EMBL: M93012; . NOT\_ANNOTATED\_CDS.  
 DR HSSP: Q99075; 1XDT.  
 DR InterPro: IPR000561; EGF-like.  
 DR Pfam: PF00008; EGF; 1.  
 DR SMART: SM00181; EGF; 1.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 DR Growth factor; Heparin-binding; Signal; EGF-like domain;  
 KW Transmembrane; Glycoprotein; Receptor.  
 FT SIGNAL 1 19  
 FT PROPEP 20 62  
 FT CHAIN 63 148  
 FT PROPEP 149 208  
 FT DOMAIN 20 160  
 FT TRANSMEM 161 184  
 FT DOMAIN 185 208  
 FT DOMAIN 104 144  
 FT CARBOHYD 75 75  
 FT CARBOHYD 85 85  
 FT DISULFID 108 121  
 FT DISULFID 116 132  
 FT DISULFID 134 143  
 FT SEQUENCE 208 AA; 22985 MW; 8D108289A0485AE9 CRC64;

Query Match 1.9%; Score 7; DB 1; Length 208;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 15 LTPSVVL 21  
 |||||  
 Db 3 LTPSVVL 9

RESULT 12  
 HBGF\_HUMAN STANDARD; PRT; 208 AA.  
 ID HBGF\_HUMAN 099075;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Heparin-binding EGF-like growth factor precursor (HB-EGF) (HBEGF)  
 DE (Diphtheria toxin receptor) (DT-R).  
 GN DTR OR HEGFL.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NC NCBL\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. AND SEQUENCE OF 73-93.  
 RX TISSUE-Macrophage;  
 RX MEDLINE=91157008; PubMed=1840698;  
 RA Hgashiyama S., Abraham J.A., Miller J., Fiddes J.C., Klagsbrun M.;  
 RT "A heparin-binding growth factor secreted by macrophage-like cells  
 that is related to EGF.";  
 RL Science 251:936-939(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX Klemmer W., Bondoc M., Cheng J., Connolly K.S., Gunning K.M.,  
 RA Kadner K., Miguel T., Miller C., Piltuck S., Pollard M., Rojeski H.,  
 RA Subramanian S., Martin C.H.;  
 RL Submitted (Apr-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 63-141 AND 143-148, AND CARBOHYDRATE-LINKAGE SITES.  
 RX TISSUE-Histiocytic lymphoma;  
 RX MEDLINE=92210596; PubMed=1556128;  
 RA Hgashiyama S., Lau K., Besner G.E., Abraham J.A., Klagsbrun M.;

RT "Structure of heparin-binding EGF-like growth factor. Multiple forms,  
RT primary structure, and glycosylation of the mature protein."  
RL J. Biol. Chem. 267:6205-6212(1992).  
RN [4]  
RP TOXIN-BINDING DOMAIN  
RX MEDLINE-95138082; PubMed-7836353;  
RA Mitamura T., Higashiyama S., Taniguchi N., Klagsbrun M., Mekada E.;  
RT "Diphtheria toxin binds to the epidermal growth factor (EGF)-like  
RT domain of human heparin-binding EGF-like growth factor/diphtheria  
RT toxin receptor and inhibits specifically its mitogenic activity.";  
RL J. Biol. Chem. 270:1015-1019(1995).  
RN [5]  
RP X-RAY CRYSTALLOGRAPHY (2.65 ANGSTROMS) OF 73-147 IN COMPLEX WITH TOX.  
RX MEDLINE-98324089; PubMed-9659904;  
RA Louie G.V., Yang W., Bowman M.E., Choe S.;  
RT "Crystal structure of the complex of diphtheria toxin with an  
RT extracellular fragment of its receptor."  
RL Mol. Cell 1:67-78(1997).  
CC -1- FUNCTION: MAY BE INVOLVED IN MACROPHAGE-MEDIATED CELLULAR  
CC PROLIFERATION. IT IS MITOGENIC FOR FIBROBLASTS AND SMOOTH MUSCLE  
CC BUT NOT ENDOTHELIAL CELLS. IT IS ABLE TO BIND EGF RECEPTORS WITH  
CC HIGHER AFFINITY THAN EGF ITSELF AND IS A FAR MORE POTENT MITOGEN  
CC FOR SMOOTH MUSCLE CELLS THAN EGF.  
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MATURE HB-EGF IS  
CC RELEASED INTO THE EXTRACELLULAR SPACE AND PROBABLY BINDS TO A  
CC RECEPTOR.  
CC -1- PTM: SEVERAL N-TERMINI HAVE BEEN IDENTIFIED BY DIRECT SEQUENCING.  
CC THE FORMS WITH N-TERMINI 63, 73 AND 74 HAVE BEEN TESTED AND FOUND  
CC TO BE BIOLOGICALLY ACTIVE.  
CC -1- PTM: O-GLYCAN ATTACHMENT SITES WERE DETERMINED BY EDMAN  
CC DEGRADATION, O-GLYCANASE DIGEST SUGGESTS MUCIN-TYPE GLYCOSYLATION  
CC (DONE IN HB-EGF PURIFIED FROM HISTIOCYTIC LYMPHOMA CELL LINE  
CC U-937).  
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
CC or send an email to [license@sdb-sib.ch](mailto:license@sdb-sib.ch)).  
CC  
CC EMBL: M60278; AAA35956.1; -  
CC EMBL: AC004634; AAC15470.1; -  
CC PIR: A38432; A38432  
CC PDB: 1XDT; 25-FEB-98.  
CC MIM: 126150; -  
CC InterPro: IPR000561; EGF-like.  
CC Pfam: PF00008; EGF; 1.  
CC SMART: SM00181; EGF; 1.  
CC PROSITE: PS00022; EGF\_1; 1.  
CC PROSITE: PS01186; EGF\_2; 1.  
CC Growth factor; Heparin-binding; Signal; EGF-like domain;  
CC Transmembrane; Glycoprotein; Receptor; 3D-structure.  
CC SIGNAL 1 19  
CC PROPEP 60 62 OR 72, OR 73, OR 76, OR 81.  
CC CHAIN 149 208 C-TERMINAL (POTENTIAL).  
CC PROPEP 149 208 C-TERMINAL (POTENTIAL).  
CC DOMAIN 20 160 EXTRACELLULAR (POTENTIAL).  
CC TRANSMEM 161 184 CYTOPLASMIC (POTENTIAL).  
CC DOMAIN 185 208 POTENTIAL.  
CC DOMAIN 104 144 EGF-LIKE.  
CC DISULFID 108 121 BY SIMILARITY.  
CC DISULFID 116 132 BY SIMILARITY.  
CC CARBOHYD 134 143 BY SIMILARITY.  
CC CARBOHYD 75 75 O-LINKED (GLNAC. . .).  
CC CARBOHYD 85 85 O-LINKED (GLNAC. . .).  
CC SEQUENCE 208 AA; 23067 MW; 2C43C9D1D8291B51 CRC64;

Query Match 1.9%; Score 7; DB 1; Length 208;  
Best Local Similarity 100.0%; Pred. No. 21;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 15 LKPSVVL 21  
Db 3 LKPSVVL 9  
RESULT 13  
HBGF\_PIG  
ID HBGF\_PIG STANDARD; PRT; 208 AA.  
AC 001560;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Heparin-binding EGF-like growth factor precursor (HB-EGF) (HBEGF).  
GN DTR OR HBGL.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Granulata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
RN NCB1\_TaxID-9823;  
RP SEQUENCE FROM N.A.  
RA Pascall J.C.;  
RL Submitted (NOV-1997) to the EMBL/Genbank/DBJ databases.  
RN [2]  
RP SEQUENCE OF 99-182 FROM N.A.  
RC STRAIN-LARGE WHITE; TISSUE-Heart;  
RX MEDLINE-93075016; PubMed-1445231;  
RA Vaughan T.J., Pascall J.C., Brown K.D.;  
RT "Tissue distribution of mRNA for heparin-binding epidermal growth  
RT factor";  
RL Biochem. J. 287:681-684(1992).  
CC -1- FUNCTION: MAY BE INVOLVED IN MACROPHAGE-MEDIATED CELLULAR  
CC PROLIFERATION. IT IS MITOGENIC FOR FIBROBLASTS AND SMOOTH MUSCLE  
CC BUT NOT ENDOTHELIAL CELLS. IT IS ABLE TO BIND EGF RECEPTORS WITH  
CC HIGHER AFFINITY THAN EGF ITSELF AND IS A FAR MORE POTENT MITOGEN  
CC FOR SMOOTH MUSCLE CELLS THAN EGF (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MATURE HB-EGF IS  
CC RELEASED INTO THE EXTRACELLULAR SPACE AND PROBABLY BINDS TO A  
CC RECEPTOR (BY SIMILARITY).  
CC -1- TISSUE SPECIFICITY: MACROPHAGE, MIDBRAIN, CEREBELLUM,  
CC HYPOTHALAMUS, CEREBRAL CORTEX, BUDBOOTHRETHAL GLAND, LUNG, HEART  
CC VENTRICLE, KIDNEY, SKIN, PROSTATE, SEMINAL VESICLE, TESTIS; AT LOW  
CC LEVELS IN LYMPH NODE, THYMUS, SPLEEN; NOT DETECTED IN PITUITARY,  
CC OLFACORY BULB, THYROID, DUODENUM, PANCREAS, LIVER, SUBMAXILLARY  
CC GLAND.  
CC -1- PTM: O-GLYCOSYLATED (PROBABLE).  
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
CC or send an email to [license@sdb-sib.ch](mailto:license@sdb-sib.ch)).  
CC  
CC EMBL: Y15731; CAA75740.1; -  
CC EMBL: X67295; CAA47709.1; -  
CC HSSP: Q99075; 1XDT.  
CC InterPro: IPR000561; EGF-like.  
CC InterPro: IPR01336; EGF\_1.  
CC Pfam: PF00008; EGF; 1.  
CC PRINTS: PRO0009; EGFTEF.  
CC SMART: SM00181; EGF; 1.  
CC PROSITE: PS00022; EGF\_1; 1.  
CC PROSITE: PS01186; EGF\_2; 1.  
CC Signal; Growth factor; Heparin-binding; EGF-like domain;  
CC Transmembrane; Glycoprotein.  
CC SIGNAL 1 23  
CC PROPEP 24 62 BY SIMILARITY.  
CC CHAIN 63 148 HEPARIN-BINDING EGF-LIKE GROWTH FACTOR.  
CC PROPEP 149 208 C-TERMINAL (POTENTIAL).

FT DOMAIN 24 161 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 162 182 POTENTIAL.  
 FT TRANSMEM 183 208 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 104 144 EGF-LIKE.  
 FT DISULFID 108 121 BY SIMILARITY.  
 FT DISULFID 116 132 BY SIMILARITY.  
 FT DISULFID 134 143 BY SIMILARITY.  
 FT CARBOHYD 85 85 O-LINKED (GALNAC. . .) (BY SIMILARITY).  
 SQ SEQUENCE 208 AA: 22986 MW: 0A7DA97AE30C8967 CRC64;

Query Match 1.9%; Score 7; DB 1; Length 208;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 15 LPSVYL 21  
 |||||  
 DB 3 LPSVYL 9

RESULT 14  
 HBGF\_RAT  
 ID HBGF\_RAT STANDARD: PRT: 208 AA.  
 AC 006175;

DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Heparin-binding EGF-like growth factor precursor (HB-EGF) (HBEGF).  
 GN DTR OR HBGF.

OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 ON NCBI\_TaxID=10116;

RP SEQUENCE FROM N.A.  
 RC STRAIN-SPRAGUE-DAWLEY; TISSUE-Macrophage;  
 RX MEDLINE=93135756; PubMed=7678488;

RA Abraham J.A., Damm D., Najdard A., Miller J., Klagsbrun M.,  
 RA Ezekowitz R.A.B.;

RT "Heparin-binding EGF-like growth factor: characterization of rat and  
 RT mouse cDNA clones, protein domain conservation across species, and  
 RT transcript expression in tissues.";

RL Biochem. Biophys. Res. Commun. 190:125-133(1993).  
 CC -1- FUNCTION: MAY BE INVOLVED IN MACROPHAGE-MEDIATED CELLULAR  
 CC PROLIFERATION. IT IS MITOGENIC FOR FIBROBLASTS AND SMOOTH MUSCLE  
 CC BUT NOT ENOTHELIAL CELLS. IT IS ABLE TO BIND EGF RECEPTORS WITH  
 CC HIGHER AFFINITY THAN EGF ITSELF AND IS A FAR MORE POTENT MITOGEN  
 CC FOR SMOOTH MUSCLE CELLS THAN EGF.

CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MATURE HB-EGF IS  
 CC RELEASED INTO THE EXTRACELLULAR SPACE AND PROBABLY BINDS TO A  
 CC RECEPTOR.  
 CC -1- TISSUE SPECIFICITY: MOST ABUNDANT IN SKELETAL MUSCLE, LUNG, SPLEEN  
 CC BRAIN AND HEART.

CC -1- PTM: O-GLYCOSYLATED (BY SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: L05489; AAA81780.1; .  
 CC PIR: JC1409; JC1409.

DR HSSP: Q99075; 1XDT.  
 DR InterPro: IPR000561; EGF-like.

DR Pfam: PF00008; EGF\_1.

DR SMART: SM00181; EGF\_1.

DR PROSITE: PS00022; EGF\_1; 1.

DR PROSITE: PS01186; EGF\_2; 1.

KW Signal; Growth factor; Heparin-binding; EGF-like domain;

KW Transmembrane; Glycoprotein.  
 FT SIGNAL 1 23 POTENTIAL.  
 FT PROPEP 24 62 BY SIMILARITY.  
 FT CHAIN 63 148 HEPARIN-BINDING EGF-LIKE GROWTH FACTOR.  
 FT PROPEP 149 208 C-TERMINAL (POTENTIAL).  
 FT DOMAIN 24 160 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 161 184 POTENTIAL.  
 FT TRANSMEM 185 208 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 104 144 EGF-LIKE.  
 FT DISULFID 108 121 BY SIMILARITY.  
 FT DISULFID 116 132 BY SIMILARITY.  
 FT DISULFID 134 143 BY SIMILARITY.  
 FT CARBOHYD 85 85 O-LINKED (GALNAC. . .) (BY SIMILARITY).  
 SQ SEQUENCE 208 AA: 22843 MW: DDBD045E116D064C CRC64;

Query Match 1.9%; Score 7; DB 1; Length 208;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 15 LPSVYL 21  
 |||||  
 DB 3 LPSVYL 9

RESULT 15  
 HXB4\_MOUSE  
 ID HXB4\_MOUSE STANDARD: PRT: 250 AA.  
 AC P10284;

DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-MAR-1989 (Rel. 10, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Homeobox protein Hox-B4 (Hox-2.6).  
 GN HOXB4 OR HOXB-4 OR HOX-2.6.

OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 ON NCBI\_TaxID=10090;

RP SEQUENCE FROM N.A.  
 RC MEDLINE=89091992; PubMed=2463210;

RA Graham A., Papalopulu N., Lorimer J., McVey J.H., Tuddenham E.G.D.,  
 RA Krutlauf R.;

RT "Characterization of a murine homeo box gene, Hox-2.6, related to the  
 RT Drosophila Deformed gene.";  
 RL Genes Dev. 2:1424-1438(1988).

CC -1- FUNCTION: SEQUENCE-SPECIFIC TRANSCRIPTION FACTOR WHICH IS PART OF  
 CC A DEVELOPMENTAL REGULATORY SYSTEM THAT PROVIDES CELLS WITH  
 CC SPECIFIC POSITIONAL IDENTITIES ON THE ANTERIOR-POSTERIOR AXIS.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- SIMILARITY: BELONGS TO THE ANTP FAMILY OF HOMEBOX PROTEINS.  
 CC "DEFORMED" SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: M36654; AAA37848.1; .  
 CC PIR: A31757; A31757.

DR HSSP: P02833; 9ANT.

DR TRANSFAC: T01728; .

DR MGD: MGI:96185; Hoxb4

DR InterPro: IPR001827; Antennapedia.

DR InterPro: IPR001356; Homeobox.

DR Pfam: PF00046; homeobox; 1.

DR PRINTS: PR00025; ANTENNAPEDIA.

DR PRINTS: PR00024; HOMEBOX.

DR SMART: SM00389; HOX; 1.

DR PROSITE: PS00027; HOMEBOX\_1; 1.

DR PROSITE: PS00032; ANTENNAPEDIA; 1.  
DR PROSITE: PS00071; HOMEOBOX\_2; 1.  
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;  
KW Transcription regulation.  
FT DOMAIN 15 138 PRO-RICH; PART OF THE TRANSCRIPTIONAL  
FT DOMAIN 71 86 ACTIVATION DOMAIN.  
FT DOMAIN 140 145 POLY-PRO.  
FT DNA\_BIND 161 220 ANTP-TYPE HEXAPEPTIDE.  
SQ SEQUENCE 250 AA; 27519 MW; D09D477A0E585BE6 CRC64;

Query Match 1.9%; Score 7; DB 1; Length 250;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 238 GGORRES 244  
Db 40 GGORRES 46

Search completed: August 28, 2002, 11:27:19  
Job time: 512 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:26:50 ; Search time 83.88 Seconds  
(without alignments)  
1050.387 Million cell updates/sec

Title: US-09-853-526-5  
364  
Sequence: 1 MLSTLVHTYSMRLLPSVY.....YVNTWIVTILGLMTWITKA 364

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 747981 seqs, 242050750 residues

Word size : 0

Total number of hits satisfying chosen parameters: 747981

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Pending\_Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	364	100.0	364	US-10-184-648-63	Sequence 63, Appl
2	269	73.9	269	US-10-074-045-47	Sequence 47, Appl
3	252	69.2	353	US-09-629-469A-13028	Sequence 13028, A
4	9	2.5	135	US-10-184-648-52	Sequence 52, Appl
5	9	2.5	157	US-10-184-648-66	Sequence 66, Appl
6	9	2.5	240	US-09-895-913A-184	Sequence 184, Appl
7	8	2.2	236	US-10-053-853A-1538	Sequence 1538, Ap
8	2.2	469	5	US-09-791-537-5820	Sequence 5820, Ap
9	8	2.2	455	US-09-791-537-110883	Sequence 110883,
10	8	2.2	483	US-09-791-537-110642	Sequence 110642,
11	7	1.9	44	US-09-826-734A-142	Sequence 142, Appl
12	7	1.9	77	US-09-548-936C-17	Sequence 17, Appl
13	7	1.9	96	US-09-791-537-150979	Sequence 150979,
14	7	1.9	117	US-09-791-537-31675	Sequence 31675, A
15	7	1.9	126	US-09-791-537-63528	Sequence 63528, A
16	7	1.9	194	US-09-791-537-53250	Sequence 53250, A
17	7	1.9	195	US-10-164-966-9	Sequence 9, Appl
18	7	1.9	195	US-10-184-648-51	Sequence 51, Appl
19	7	1.9	195	US-10-184-648-67	Sequence 67, Appl
20	7	1.9	195	US-10-184-648-86	Sequence 86, Appl
21	7	1.9	208	PCT-US02-10824-132	Sequence 132, Appl
22	7	1.9	208	US-09-791-537-57367	Sequence 57367, A
23	7	1.9	208	US-09-791-537-99121	Sequence 99121, A
24	7	1.9	208	US-09-791-537-128069	Sequence 128069,
25	7	1.9	208	US-09-791-537-142901	Sequence 142901,
26	7	1.9	208	US-09-791-537-143774	Sequence 143774,

27	7	1.9	208	US-10-096-327-2	Sequence 2, Appl
28	7	1.9	208	US-10-179-131-7124	Sequence 7124, Ap
29	7	1.9	208	US-10-189-360-5	Sequence 5, Appl
30	7	1.9	208	US-10-189-360-7	Sequence 7, Appl
31	7	1.9	208	US-10-189-360-8	Sequence 8, Appl
32	7	1.9	208	US-10-138-158-18	Sequence 18, Appl
33	7	1.9	234	US-09-791-537-148811	Sequence 148811,
34	7	1.9	246	US-09-791-537-43519	Sequence 43519, A
35	7	1.9	250	US-09-791-537-67119	Sequence 67119, A
36	7	1.9	251	US-09-791-537-29660	Sequence 29660, A
37	7	1.9	255	US-10-153-881-22428	Sequence 22428, A
38	7	1.9	262	US-09-791-537-751	Sequence 751, Appl
39	7	1.9	263	US-10-155-881-18654	Sequence 18654, A
40	7	1.9	264	US-09-791-537-24942	Sequence 24942, A
41	7	1.9	269	US-10-053-853A-1244	Sequence 1244, Ap
42	7	1.9	270	US-10-211-364-1026	Sequence 1026, Ap
43	7	1.9	286	US-60-360-039-1733	Sequence 1733, Ap
44	7	1.9	311	US-60-360-039-22196	Sequence 22196, A
45	7	1.9	364	US-09-791-537-42339	Sequence 42339, A

ALIGNMENTS

RESULT 1  
US-10-184-648-63  
Sequence 63, Application US/10184648  
GENERAL INFORMATION:  
APPLICANT: Meyers, Rachel E.  
APPLICANT: Williamson, Mark  
APPLICANT: Tsai, Feng-Ying  
APPLICANT: Hunter, John J.  
APPLICANT: Macbeth, Kyle J.  
APPLICANT: Rudolph-Owen, Laura A.  
APPLICANT: Leiby, Kevin R.  
APPLICANT: Kapeller-Libermann, Rosana  
APPLICANT: Olandt, Peter J.  
TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE FAMILY MEMBERS AND USES THEREOF  
FILE REFERENCE: 10448-192001  
CURRENT APPLICATION NUMBER: US/10/184,648  
CURRENT FILING DATE: 2002-06-27  
PRIOR APPLICATION NUMBER: US 09/815,028  
PRIOR FILING DATE: 2001-03-22  
PRIOR APPLICATION NUMBER: PCT/US01/09358  
PRIOR FILING DATE: 2001-03-22  
PRIOR APPLICATION NUMBER: US 60/191,964  
PRIOR FILING DATE: 2000-03-24  
PRIOR APPLICATION NUMBER: US 09/801,220  
PRIOR FILING DATE: 2001-03-07  
PRIOR APPLICATION NUMBER: PCT/US01/07269  
PRIOR FILING DATE: 2001-03-07  
PRIOR APPLICATION NUMBER: US 60/187,456  
PRIOR FILING DATE: 2000-03-07  
PRIOR APPLICATION NUMBER: US 09/816,714  
PRIOR FILING DATE: 2001-03-23  
PRIOR APPLICATION NUMBER: PCT/US01/09468  
PRIOR FILING DATE: 2001-03-23  
PRIOR APPLICATION NUMBER: US 60/191,865  
PRIOR FILING DATE: 2000-03-24  
PRIOR APPLICATION NUMBER: US 09/844,948  
PRIOR FILING DATE: 2001-04-27  
PRIOR APPLICATION NUMBER: PCT/US01/13805  
PRIOR FILING DATE: 2001-04-27  
PRIOR APPLICATION NUMBER: US 60/200,604  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: US 09/861,164  
PRIOR FILING DATE: 2001-05-18  
PRIOR APPLICATION NUMBER: PCT/US01/16292  
PRIOR FILING DATE: 2001-05-18  
PRIOR APPLICATION NUMBER: US 60/205,408  
PRIOR FILING DATE: 2000-05-19  
PRIOR APPLICATION NUMBER: US 09/883,060  
PRIOR FILING DATE: 2001-06-15

```

; PRIOR APPLICATION NUMBER: PCT/US01/19138
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: JS 60/212,079
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/962,678
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: PCT/US01/29963
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/235,044
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 09/973,457
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: US 60/238,849
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 10/072,285
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: PCT/US02/03736
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/267,494
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: US 09/817,910
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: PCT/US01/09633
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/192,092
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: US 09/842,528
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: PCT/US01/40607
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 60/199,500
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/882,836
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19543
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/211,730
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/882,872
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19153
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 50/212,077
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-184-648-63
```

```

Query Match      100.0%; Score 364; DB 6; Length 364;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 MLSTLVHTYSMKRYLLPSVLLGTAPTYVLAMGVWRLLSAFLPARFYQALDDRLCYCYOS 60
DB 1 MLSTLVHTYSMKRYLLPSVLLGTAPTYVLAMGVWRLLSAFLPARFYQALDDRLCYCYOS 50
QY 61 MVLFFENYTGVOITLYGDLPRKN:ENITTYLANHSTVDWIVADILAIROMALGHVRYLK 120
DB 61 MVLFFENYTGVOITLYGDLPRKN:ENITTYLANHSTVDWIVADILAIROMALGHVRYLK 120
QY 121 EGKMLPLVGCYFAOHGCIYVRSKAFNEKEMKNKLSQSYVDAGTPLYLVFFPESTRNPE 180
DB 121 EGKMLPLVGCYFAOHGCIYVRSKAFNEKEMKNKLSQSYVDAGTPLYLVFFPESTRNPE 180
QY 181 QTVLSASQAFAGRLAVLKHVLPRIKATHVAFDCMKRYLDAIYDVTVYVGGKDDGGQ 240
DB 181 QTVLSASQAFAGRLAVLKHVLPRIKATHVAFDCMKRYLDAIYDVTVYVGGKDDGGQ 240
```

```

QY 241 RRESPTWTEFLCKECPKIHIIHIDRDKDVPEDQEMRRMLHREFEIKDKMLIEFESPD 300
DB 241 RRESPTWTEFLCKECPKIHIIHIDRDKDVPEDQEMRRMLHREFEIKDKMLIEFESPD 300
QY 301 PERKRPFGKSVNSKLSIKKTPSMLISGLTAGMLTADGRKLYVTWTYGTLLGLMW 360
DB 301 PERKRPFGKSVNSKLSIKKTPSMLISGLTAGMLTADGRKLYVTWTYGTLLGLMW 360
QY 361 TIRA 364
DB 361 TIRA 364
```

## RESULT 2

```

US-10-074-045-47
; Sequence 47, Application US/10074045
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT221C1
; CURRENT APPLICATION NUMBER: US/10/074,045
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 47
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-074-045-47
```

```

Query Match      73.9%; Score 269; DB 6; Length 269;
Best Local Similarity 100.0%; Pred. No. 1.5e-267;
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 96 TVDMIVADILAIROMALGHVRYLKESGLKMLPLYGCTFAOHGCIYVRSKAFNEKEMKNK 155
DB 1 TVDMIVADILAIROMALGHVRYLKESGLKMLPLYGCTFAOHGCIYVRSKAFNEKEMKNK 155
QY 156 LQSYVDAGTPLYLVFFPESTRNPEQTKVLSASQAFAGRLAVLKHVLPRIKATHVAF 215
DB 61 LQSYVDAGTPLYLVFFPESTRNPEQTKVLSASQAFAGRLAVLKHVLPRIKATHVAF 120
QY 216 DCKKNYLDIAYDVTVYVGGKDDGQRRSEPTWTEFLCKECPKIHIIHIDRDKDVEEOE 275
DB 121 DCKKNYLDIAYDVTVYVGGKDDGQRRSEPTWTEFLCKECPKIHIIHIDRDKDVEEOE 180
QY 276 HMRRLHREFEIKDKMLIEFESPDPERKRPFGKSVNSKLSIKKTPSMLISGLTAGM 335
DB 181 HMRRLHREFEIKDKMLIEFESPDPERKRPFGKSVNSKLSIKKTPSMLISGLTAGM 240
QY 336 LMTDAGRKLIVNTWYGTLLGLMWITIKA 364
DB 241 LMTDAGRKLIVNTWYGTLLGLMWITIKA 269
```

## RESULT 3

```

US-09-629-469A-13028
; Sequence 13028, Application US/09629469A
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: SAITO, KAOBU
; APPLICANT: YAMAMOTO, JUNICHI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: WAKAMATSU, AI
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: OTSUKI, TETSUJI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE
```

```
FILE REFERENCE: 084335/0123
CURRENT APPLICATION NUMBER: US/09/629,469A
CURRENT FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: JP 1999-248036
PRIOR FILING DATE: 1999-07-29
PRIOR APPLICATION NUMBER: JP 1999-300253
PRIOR FILING DATE: 1999-08-27
PRIOR APPLICATION NUMBER: JP 2000-118776
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: JP 2000-183767
PRIOR FILING DATE: 2000-05-02
PRIOR APPLICATION NUMBER: JP 2000-241899
PRIOR FILING DATE: 2000-06-09
PRIOR APPLICATION NUMBER: 60/159,590
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 60/183,322
PRIOR FILING DATE: 2000-02-17
NUMBER OF SEQ ID NOS: 19025
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 13028
LENGTH: 353
TYPE: PRT
ORGANISM: Homo sapiens
US-09-629-469A-13028
```

```
Query Match          69.2%  Score 252; DB 5; Length 353;
Best Local Similarity 99.7%  Pred. No. 5.7e-250;
Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 12 MRYLPSVVLGTAPTYVYLVAMGWRLSAPLPAFYQALDDRCYCYQSQVLEFFENYTG 71
Db 1 MRYLPSVVLGTAPTYVYLVAMGWRLSAPLPAFYQALDDRCYCYQSQVLEFFENYTG 60
QY 72 VQILLGYDLPRKNENIYIYLANHOSYDWMYADILAIQNALGHVRYLKEGLKWLPLYGC 131
Db 61 VQILLGYDLPRKNENIYIYLANHOSYDWMYADILAIQNALGHVRYLKEGLKWLPLYGC 120
QY 132 YFADHGGIYVYKSAKFNEMKRNKLOSVDAGTPTMYIVTPECTRYVPEQTKYLSAQAF 191
Db 121 YFADHGGIYVYKSAKFNEMKRNKLOSVDAGTPTMYIVTPECTRYVPEQTKYLSAQAF 180
QY 192 AAGRGVAVLKHVLPRIKATHVAFDCMKNYLDATYDVTYVYEGKDGQRRSEPTWTEFL 251
Db 181 AAGRGVAVLKHVLPRIKATHVAFDCMKNYLDATYDVTYVYEGKDGQRRSEPTWTEFL 240
QY 252 CKECPRTHIHIDRIDKKDVEEQQEHMRWMLHERFEIKDKMLIEFYESPDERRRRPFPGKS 311
Db 241 CKECPRTHIHIDRIDKKDVEEQQEHMRWMLHERFEIKDKMLIEFYESPDERRRRPFPGKS 300
QY 312 VNSKLSIKKTLPMSLLISGLTAGMLMTDAGRKIYVNTWYIGTLGLCMTWITKA 364
Db 301 VNSKLSIKKTLPMSLLISGLTAGMLMTDAGRKIYVNTWYIGTLGLCMTWITKA 353
```

```
RESULT 4
US-10-184-648-52
Sequence 52, Application US/10184648
GENERAL INFORMATION:
APPLICANT: Meyers, Rachel E.
APPLICANT: Williamson, Mark
APPLICANT: Tsai, Feng-Ying
APPLICANT: Hunter, John J.
APPLICANT: Macbeth, Kyle J.
APPLICANT: Rudolph-Owen, Laura A.
APPLICANT: Leiby, Kevin R.
APPLICANT: Kapeller-Libermann, Rosana
APPLICANT: Olandt, Peter J.
TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE FAMILY MEMBERS AND USES THEREOF
FILE REFERENCE: 10448-192001
CURRENT APPLICATION NUMBER: US/10/184,648
CURRENT FILING DATE: 2002-06-27
PRIOR APPLICATION NUMBER: US 09/815,028
```

```
PRIOR FILING DATE: 2001-03-22
PRIOR APPLICATION NUMBER: PCT/US01/09358
PRIOR FILING DATE: 2001-03-22
PRIOR APPLICATION NUMBER: US 60/191,964
PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: US 09/801,220
PRIOR FILING DATE: 2001-03-07
PRIOR APPLICATION NUMBER: PCT/US01/07269
PRIOR FILING DATE: 2001-03-07
PRIOR APPLICATION NUMBER: US 60/187,456
PRIOR FILING DATE: 2000-03-07
PRIOR APPLICATION NUMBER: US 09/816,714
PRIOR FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: PCT/US01/09468
PRIOR FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: US 60/191,865
PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: US 09/844,948
PRIOR FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: PCT/US01/13805
PRIOR FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 60/200,604
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 09/861,164
PRIOR FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: PCT/US01/16292
PRIOR FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: US 60/205,408
PRIOR FILING DATE: 2000-05-19
PRIOR APPLICATION NUMBER: US 09/883,060
PRIOR FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: PCT/US01/19138
PRIOR FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: US 60/212,079
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: US 09/962,678
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: PCT/US01/29963
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US 60/235,044
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US 09/973,457
PRIOR FILING DATE: 2001-10-09
PRIOR APPLICATION NUMBER: US 60/238,849
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: US 10/072,285
PRIOR FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: PCT/US02/03736
PRIOR FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: US 60/267,494
PRIOR FILING DATE: 2001-02-08
PRIOR APPLICATION NUMBER: US 09/817,910
PRIOR FILING DATE: 2001-03-26
PRIOR APPLICATION NUMBER: PCT/US01/09633
PRIOR FILING DATE: 2001-03-26
PRIOR APPLICATION NUMBER: US 60/192,092
PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: US 09/842,528
PRIOR FILING DATE: 2001-04-25
PRIOR APPLICATION NUMBER: PCT/US01/40607
PRIOR FILING DATE: 2001-04-25
PRIOR APPLICATION NUMBER: US 60/199,500
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: US 09/882,836
PRIOR FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: PCT/US01/19543
PRIOR FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: US 60/211,730
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: US 09/882,872
PRIOR FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: PCT/US01/19153
PRIOR FILING DATE: 2001-06-15
```

;; PRIOR APPLICATION NUMBER: US 60/212,077  
;; PRIOR FILING DATE: 2000-06-15  
;; NUMBER OF SEQ ID NOS: 90  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 52  
;; LENGTH: 135  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: consensus sequence  
US-10-184-648-52

Query Match 2.5%; Score 9; DB 6; Length 135;  
Best Local Similarity 100.0%; Pred. No. 0.71;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 168 LVIFPEGTR 176  
Db 85 LVIFPEGTR 93

RESULT 5  
US-10-184-648-66  
;; Sequence 66, Application US/101846648  
;; GENERAL INFORMATION:  
;; APPLICANT: Meyers, Rache, E.  
;; APPLICANT: Williamson, Mark  
;; APPLICANT: Tsai, Fong-Ying  
;; APPLICANT: Hunter, John J.  
;; APPLICANT: Macbeth, Kyle J.  
;; APPLICANT: Rudolph-Owen, Laura A.  
;; APPLICANT: Leiby, Kevin R.  
;; APPLICANT: Kapeller-Libermann, Rosana  
;; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE FAMILY MEMBERS AND USES THEREOF  
;; FILE REFERENCE: 1048-192001  
;; CURRENT APPLICATION NUMBER: US/10/184,648  
;; PRIOR FILING DATE: 2002-06-27  
;; PRIOR APPLICATION NUMBER: US 09/815,028  
;; PRIOR FILING DATE: 2001-03-22  
;; PRIOR APPLICATION NUMBER: PCT/US01/09358  
;; PRIOR FILING DATE: 2001-03-22  
;; PRIOR APPLICATION NUMBER: US 60/191,964  
;; PRIOR FILING DATE: 2000-03-24  
;; PRIOR APPLICATION NUMBER: US 09/801,220  
;; PRIOR FILING DATE: 2001-03-07  
;; PRIOR APPLICATION NUMBER: PCT/US01/07269  
;; PRIOR FILING DATE: 2001-03-07  
;; PRIOR APPLICATION NUMBER: US 60/187,456  
;; PRIOR FILING DATE: 2000-03-07  
;; PRIOR APPLICATION NUMBER: US 09/816,714  
;; PRIOR FILING DATE: 2001-03-23  
;; PRIOR APPLICATION NUMBER: PCT/US01/09468  
;; PRIOR FILING DATE: 2001-03-23  
;; PRIOR APPLICATION NUMBER: US 60/191,865  
;; PRIOR FILING DATE: 2000-03-24  
;; PRIOR APPLICATION NUMBER: US 09/844,948  
;; PRIOR FILING DATE: 2001-04-27  
;; PRIOR APPLICATION NUMBER: PCT/US01/13805  
;; PRIOR FILING DATE: 2001-04-27  
;; PRIOR APPLICATION NUMBER: US 60/200,604  
;; PRIOR FILING DATE: 2000-04-28  
;; PRIOR APPLICATION NUMBER: US 09/861,164  
;; PRIOR FILING DATE: 2001-05-18  
;; PRIOR APPLICATION NUMBER: PCT/US01/16292  
;; PRIOR FILING DATE: 2001-05-18  
;; PRIOR APPLICATION NUMBER: US 60/205,408  
;; PRIOR FILING DATE: 2000-05-19  
;; PRIOR APPLICATION NUMBER: US 09/883,060  
;; PRIOR FILING DATE: 2001-06-15  
;; PRIOR APPLICATION NUMBER: PCT/US01/19138  
;; PRIOR FILING DATE: 2001-06-15

;; PRIOR APPLICATION NUMBER: US 60/212,079  
;; PRIOR FILING DATE: 2000-06-15  
;; PRIOR APPLICATION NUMBER: US 09/962,678  
;; PRIOR FILING DATE: 2001-09-25  
;; PRIOR APPLICATION NUMBER: PCT/US01/29963  
;; PRIOR FILING DATE: 2001-09-25  
;; PRIOR APPLICATION NUMBER: US 60/235,044  
;; PRIOR FILING DATE: 2000-09-25  
;; PRIOR APPLICATION NUMBER: US 09/973,457  
;; PRIOR FILING DATE: 2001-10-09  
;; PRIOR APPLICATION NUMBER: US 60/238,849  
;; PRIOR FILING DATE: 2000-10-06  
;; PRIOR APPLICATION NUMBER: US 10/072,285  
;; PRIOR FILING DATE: 2002-02-08  
;; PRIOR APPLICATION NUMBER: PCT/US02/03736  
;; PRIOR FILING DATE: 2002-02-08  
;; PRIOR APPLICATION NUMBER: US 60/267,494  
;; PRIOR FILING DATE: 2001-02-08  
;; PRIOR APPLICATION NUMBER: US 09/817,910  
;; PRIOR FILING DATE: 2001-03-26  
;; PRIOR APPLICATION NUMBER: PCT/US01/09633  
;; PRIOR FILING DATE: 2001-03-26  
;; PRIOR APPLICATION NUMBER: US 60/192,092  
;; PRIOR FILING DATE: 2000-03-24  
;; PRIOR APPLICATION NUMBER: US 09/842,528  
;; PRIOR FILING DATE: 2001-04-25  
;; PRIOR APPLICATION NUMBER: PCT/US01/40607  
;; PRIOR FILING DATE: 2001-04-25  
;; PRIOR APPLICATION NUMBER: US 60/199,500  
;; PRIOR FILING DATE: 2000-04-25  
;; PRIOR APPLICATION NUMBER: US 09/882,836  
;; PRIOR FILING DATE: 2001-06-15  
;; PRIOR APPLICATION NUMBER: PCT/US01/19543  
;; PRIOR FILING DATE: 2001-06-15  
;; PRIOR APPLICATION NUMBER: US 60/211,730  
;; PRIOR FILING DATE: 2000-06-15  
;; PRIOR APPLICATION NUMBER: US 09/882,872  
;; PRIOR FILING DATE: 2001-06-15  
;; PRIOR APPLICATION NUMBER: PCT/US01/19153  
;; PRIOR FILING DATE: 2001-06-15  
;; PRIOR APPLICATION NUMBER: US 60/212,077  
;; PRIOR FILING DATE: 2000-06-15  
;; NUMBER OF SEQ ID NOS: 90  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 66  
;; LENGTH: 157  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: consensus sequence  
US-10-184-648-66

Query Match 2.5%; Score 9; DB 6; Length 157;  
Best Local Similarity 100.0%; Pred. No. 0.81;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 222 LDATIDVTY 230  
Db 4 LDATIDVTY 12

RESULT 6  
US-09-895-913A-184  
;; Sequence 184, Application US/09895913A  
;; GENERAL INFORMATION:  
;; APPLICANT: Kleantous, Harold  
;; APPLICANT: Al-Garawi, Amal  
;; APPLICANT: Miller, Charles  
;; APPLICANT: Tomb, Jean Francois  
;; APPLICANT: Oomen, Raymond P.  
;; TITLE OF INVENTION: Identification of Polynucleotides  
;; TITLE OF INVENTION: Encoding Novel Helicobacter Polypeptides in the Helicobacter

```
; TITLE OF INVENTION: Genome
; FILE REFERENCE: 06132/043002
; CURRENT APPLICATION NUMBER: US/09/895,913A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 08/881,227
; PRIOR FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 184
; LENGTH: 240
; TYPE: PRF
; ORGANISM: Helicobacter pylori
US-09-895-913A-184
```

```
Query Match          2.5%; Score 9; DB 5; Length 240;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 168 LVIFPESTR 176
    |||||
Db 142 LVIFPESTR 150
```

```
RESULT 7
US-10-053-853A-1538
; Sequence 1538, Application US/10053853A
; GENERAL INFORMATION:
; APPLICANT: HAYASHI, Hideo
; APPLICANT: SHINGAWA, Hideo
; APPLICANT: MAKINO, Kozo
; APPLICANT: HAYASHI, Tetsuya
; APPLICANT: OHNISHI, Makoto
; APPLICANT: HATTORI, Masahira
; APPLICANT: KUROKAWA, Ken
; TITLE OF INVENTION: Polynucleotide molecules and polypeptides specific to Enterohemorrhagic E. coli O157:H7 and use thereof
; FILE REFERENCE: 2002-0060A/MNC/01704
; CURRENT APPLICATION NUMBER: US/10/053,853A
; CURRENT FILING DATE: 2002-05-28
; PRIOR APPLICATION NUMBER: JP2001-112010
; PRIOR FILING DATE: 2001-01-24
; NUMBER OF SEQ ID NOS: 1866
; SEQ ID NO 1538
; LENGTH: 236
; TYPE: PRF
; ORGANISM: Escherichia coli O157:H7
US-10-053-853A-1538
```

```
Query Match          2.2%; Score 8; DB 6; Length 236;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 204 LTPRIKAT 211
    |||||
Db 227 LTPRIKAT 234
```

```
RESULT 8
US-09-791-537-5820
; Sequence 5820, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Dabe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
```

```
; SEQ ID NO 5820
; LENGTH: 262
; TYPE: PRF
; ORGANISM: Chlamydia trachomatis
US-09-791-537-5820
```

```
Query Match          2.2%; Score 8; DB 5; Length 262;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 202 HVTPESTR 209
    |||||
Db 202 HVTPESTR 209
```

```
RESULT 9
US-09-791-537-110883
; Sequence 110883, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Dabe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 110883
; LENGTH: 459
; TYPE: PRF
; ORGANISM: Bacillus subtilis
US-09-791-537-110883
```

```
Query Match          2.2%; Score 8; DB 5; Length 459;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 178 NPEQTKVL 185
    |||||
Db 292 NPEQTKVL 299
```

```
RESULT 10
US-09-791-537-110642
; Sequence 110642, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Dabe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 110642
; LENGTH: 483
; TYPE: PRF
; ORGANISM: Pisum sativum
US-09-791-537-110642
```

```
Query Match          2.2%; Score 8; DB 5; Length 483;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 162 AGTPMYLV 169
    |||||
```

Db 184 AGTPMYLV 191

RESULT 11

US-09-826-734A-142

; Sequence 142, Application US/09826734A

; GENERAL INFORMATION:

; APPLICANT: Fernandes, Nima R

; APPLICANT: Mishra, Vishnu S

; APPLICANT: Leach, Martin D

; APPLICANT: Shlmkets, Richard A

; APPLICANT: Zerhusen, Bryan D

; APPLICANT: Kekuda, Ramash

; TITLE OF INVENTION: Novel Polynucleotides And Polypeptides Encoded Thereby

; FILE REFERENCE: 15966-754

; CURRENT APPLICATION NUMBER: US/09/826,734A

; CURRENT FILING DATE: 2001-04-05

; PRIOR APPLICATION NUMBER: 60/195,576

; PRIOR FILING DATE: 2000-04-26

; NUMBER OF SEQ ID NOS: 264

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 142

; LENGTH: 44

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-826-734A-142

Query Match

Best Local Similarity 1.9%; Score 7; DB 5; Length 44;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 170 IFEPTGR 176

Db 3 IFEPTGR 9

RESULT 12

US-09-548-936C-17

; Sequence 17, Application US/09548936C

; GENERAL INFORMATION:

; APPLICANT: Gopez, Leonel Jorge

; APPLICANT: Sares, Jan

; APPLICANT: Claesson-Welsh, Lena

; APPLICANT: Heldin, Carl-Henrik

; TITLE OF INVENTION: PTPL BINDING AGENTS

; FILE REFERENCE: L0461/7084

; CURRENT APPLICATION NUMBER: US/09/548,936C

; CURRENT FILING DATE: 2000-04-13

; PRIOR APPLICATION NUMBER: US 09/100,804

; PRIOR FILING DATE: 1998-06-19

; PRIOR APPLICATION NUMBER: US 08/596,291

; PRIOR FILING DATE: 1994-09-01

; PRIOR APPLICATION NUMBER: US 08/115,573

; PRIOR FILING DATE: 1993-09-01

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: PatentIn Version 3.0

; SEQ ID NO 17

; LENGTH: 77

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-548-936C-17

Query Match

Best Local Similarity 1.9%; Score 7; DB 5; Length 77;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 136 HGGIYVK 142

Db 20 HGGIYVK 26

RESULT 13

US-09-791-537-150979

; Sequence 150979, Application US/09791537

; GENERAL INFORMATION:

; APPLICANT: Biomimix, Inc.

; APPLICANT: Debe, Derek

; APPLICANT: Danzer, Joseph

; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME

; FILE REFERENCE: 261/210

; CURRENT APPLICATION NUMBER: US/09/791,537

; CURRENT FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 153055

; SOFTWARE: PatentIn Version 3.0

; SEQ ID NO 150979

; LENGTH: 96

; TYPE: PRT

; ORGANISM: pdb 3PDZA

; FEATURE:

; NAME/KEY: misc.feature

; LOCATION: (1)..(1)

; OTHER INFORMATION: X is an unknown amino acid

US-09-791-537-150979

Query Match

Best Local Similarity 1.9%; Score 7; DB 5; Length 96;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 136 HGGIYVK 142

Db 32 HGGIYVK 38

RESULT 14

US-09-791-537-31675

; Sequence 31675, Application US/09791537

; GENERAL INFORMATION:

; APPLICANT: Biomimix, Inc.

; APPLICANT: Debe, Derek

; APPLICANT: Danzer, Joseph

; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME

; FILE REFERENCE: 261/210

; CURRENT APPLICATION NUMBER: US/09/791,537

; CURRENT FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 153055

; SOFTWARE: PatentIn Version 3.0

; SEQ ID NO 31675

; LENGTH: 117

; TYPE: PRT

; ORGANISM: Mus musculus

US-09-791-537-31675

Query Match

Best Local Similarity 1.9%; Score 7; DB 5; Length 117;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 136 HGGIYVK 142

Db 43 HGGIYVK 49

RESULT 15

US-09-791-537-63528

; Sequence 63528, Application US/09791537

; GENERAL INFORMATION:

; APPLICANT: Biomimix, Inc.

; APPLICANT: Debe, Derek

; APPLICANT: Danzer, Joseph

; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME

; FILE REFERENCE: 261/210

; CURRENT APPLICATION NUMBER: US/09/791,537

; CURRENT FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 153055

; SOFTWARE: PatentIn Version 3.0

; SEQ ID NO 31675

; LENGTH: 117

; TYPE: PRT

; ORGANISM: Mus musculus

US-09-791-537-63528

; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 63528  
; LENGTH: 126  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-791-537-63528

Query Match 1.9%; Score 7; DB 5; Length 126;  
Best Local Similarity 100.0%; Pred. No. 75;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 136 HGGIYVK 142  
Db 52 HGGIYVK 58

Search completed: August 28, 2002, 11:26:50  
Job time: 563 sec

**THIS PAGE BLANK (uspto)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:20:08 ; Search time 39.16 Seconds  
(without alignments)  
893.169 Million cell updates/sec

Title: US-09-853-526-5  
364  
Sequence: 1 MLSLVLTHTYSMRYLPSVY.....YVNTWIVGTLLGLMTWITKA 364

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283138 seqs, 96089334 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :  
1: PIR\_71:\*  
2: pir1:\*  
3: pir2:\*  
4: pir3:\*  
5: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	2.5	237	B71827	probable 1-acylgly
2	9	2.5	240	D64688	probable 1-acylgly
3	9	2.5	264	E82903	1-acyl-sn-glycerol
4	9	2.5	918	T34057	hypothetical prote
5	8	2.2	83	W8BP67	gene 18.7 protein
6	8	2.2	83	W8BP67	gene 18.7 protein
7	8	2.2	236	H90892	probable fibritial
8	8	2.2	236	B87524	1-acyl-sn-glycerol
9	8	2.2	239	A85725	probable fibritial
10	8	2.2	262	A71478	probable metal dep
11	8	2.2	282	G71334	probable lysophosp
12	8	2.2	396	S54641	probable membrane
13	8	2.2	411	C87586	metal ion efflux m
14	8	2.2	424	T49607	guanine deaminase
15	8	2.2	458	G83735	RNA methyltransfer
16	8	2.2	459	E69793	RNA methyltransfer
17	8	2.2	483	T06459	62K sucrose-bindin
18	8	2.2	611	A85000	heat shock protein
19	7	1.9	54	S58122	heparin-binding Eg
20	7	1.9	57	D86108	hypothetical prote
21	7	1.9	117	I81209	lysine phosphata
22	7	1.9	126	S58117	EGF-like growth fa
23	7	1.9	141	B90523	hypothetical prote
24	7	1.9	143	C86732	transcription regu
25	7	1.9	148	PC1136	amitophosphoribosy
26	7	1.9	194	C83854	hypothetical prote
27	7	1.9	195	A38432	heparin-binding Eg
28	7	1.9	208	A41914	diphtheria toxin re
29	7	1.9	208	A41914	diphtheria toxin re

30	7	1.9	208	1	JC1409	heparin-binding Eg
31	7	1.9	209	2	B83329	probable acyltrans
32	7	1.9	210	2	A10342	probable acyltrans
33	7	1.9	211	2	E70476	2-acylglycerophosp
34	7	1.9	225	2	H95244	conserved hypothet
35	7	1.9	226	2	E98109	conserved hypothet
36	7	1.9	230	2	E81397	probable 1-acylgly
37	7	1.9	240	2	S75162	hypothetical prote
38	7	1.9	241	2	B97019	1-acyl-sn-glycerol
39	7	1.9	246	2	T46446	hypothetical prote
40	7	1.9	247	2	A81957	1-acylglycerol-3-p
41	7	1.9	247	2	G81013	1-acyl-sn-glycerol
42	7	1.9	250	1	A31757	homeotic protein H
43	7	1.9	250	2	T27772	26S proteasome reg
44	7	1.9	250	2	E70104	1-acylglycerol-3-p
45	7	1.9	250	2	AF3384	outer membrane pro

ALIGNMENTS

RESULT 1  
B71827  
Probable 1-acylglycerol-3-phosphate O-acyltransferase (EC 2.3.1.51) - Helicobacter pylori  
C:Species: Helicobacter pylori  
A:Variate: strain J99  
C>Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 06-Oct-2000  
C:Accession: B71827  
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.  
; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.  
Nature 397, 176-180, 1999  
A>Title: Genomic sequence comparison of two unrelated isolates of the human gastric p  
A:Reference number: A71800; MUID:99120557  
A:Accession: B71827  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-237 <ARN>  
A:Cross-references: GB:AE001550; GB:AE001439; NID:g4155872; PIDN:AAD06852.1; PID:g415  
A:Experimental source: strain J99  
C:Genetics:  
A:Gene: p13C  
C:Superfamily: mouse 1-acylglycerol-3-phosphate O-acyltransferase  
C:Keywords: acyltransferase; coenzyme A

Query Match 2.5%; Score 9; DB 2; Length 237;  
Best Local Similarity 100.0%; Pred. No. 0.38;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 LVIFPEGTR 176  
DB 142 LVIFPEGTR 150

RESULT 2  
D64688  
Probable 1-acylglycerol-3-phosphate O-acyltransferase (EC 2.3.1.51) - Helicobacter pylori  
C:Species: Helicobacter pylori  
C>Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 06-Oct-2000  
C:Accession: D64688  
R:Tom, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.  
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khaliq, H.G.; Glodek, A.; McKe  
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Wattey,  
Nature 388, 539-547, 1997  
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karp, P.D.; Smith, H.O.; Fraser,  
A>Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A:Reference number: A64520; MUID:97394467  
A:Accession: D64688  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-240 <TOM>  
A:Cross-references: GB:AE000636; GB:AE000511; NID:g2314517; PIDN:AAD08393.1; PID:g231  
C:Superfamily: mouse 1-acylglycerol-3-phosphate O-acyltransferase

C;Keywords: acyltransferase; coenzyme A

Query Match 2.5%; Score 9; DB 2; Length 240;  
Best Local Similarity 100.0%; Pred. No. 0.39;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 168 LVIFPECTR 176  
Db 142 LVIFPECTR 150

RESULT 3  
E82903  
1-acyl-sn-glycerol-3-phosphate acyltransferase U0344 [imported] - Ureaplasma urealyticum  
C;Species: Ureaplasma urealyticum  
C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C;Accession: E82903  
R;Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Helner, C.R.; Chen, E.Y.; Cassell, G.H.  
submitted to GenBank, February 2000  
A;Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mit  
A;Reference number: A82870  
A;Accession: E82903  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-264 <GLA>  
A;Cross-references: GB:AE002131; GB:AF222894; NID:96899316; PIDN:AAF30753.1; GSPDB:GN001  
A;Experimental source: serovar 3; biovar 1  
C;Genetics:  
A;Gene: pLSC; U0344  
A;Genetic code: SGC3

Query Match 2.5%; Score 9; DB 2; Length 264;  
Best Local Similarity 100.0%; Pred. No. 0.42;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 168 LVIFPECTR 176  
Db 156 LVIFPECTR 164

RESULT 4  
T34057  
hypothetical protein F28B3.5 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
C;Accession: T34057  
R;Geisler, C.; Kramer, J.; Smith, A.  
submitted to the EMBL Data Library, May 1997  
A;Description: The sequence of C. elegans cosmid F28B3.  
A;Reference number: Z21469  
A;Accession: T34057  
A;Status: preliminary; translated from GB/EMBL/DBD  
A;Molecule type: DNA  
A;Residues: 1-918 <GET>  
A;Cross-references: EMBL:AF003136; PIDN:AA83636.1; GSPDB:GN00019; CESP:F28B3.5  
A;Experimental source: strain Bristol N2; clone F28B3  
C;Genetics:  
A;Gene: CESP:F28B3.5  
A;Map position: 1  
A;Introns: 85/3; 129/3; 235/3; 410/2; 482/3; 532/3; 736/3; 829/3; 850/2

Query Match 2.5%; Score 9; DB 2; Length 918;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 222 LDAIVDTV 230  
Db 762 LDAIVDTV 770

RESULT 5  
W8BP7  
gene 18.7 protein - phage T7  
C;Species: phage T7  
C;Date: 13-Jun-1983 #sequence\_revision 13-Jun-1983 #text\_change 23-Jul-1999  
C;Accession: A04413; S42337  
R;Dunn, J.C.; Thompson, K.  
submitted to the Nucleic Acid Sequence Database, September 1982  
A;Reference number: A94615  
A;Accession: A04413  
A;Molecule type: DNA  
A;Residues: 1-83 <DUN>  
R;Dunn, J.C.; Studier, F.W.  
J. Mol. Biol. 166, 477-535, 1983  
A;Title: Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7  
A;Reference number: S42283; MUID:83241725  
A;Accession: S42337  
A;Molecule type: DNA  
A;Residues: 1-83 <DUN>  
A;Cross-references: EMBL:V01146; NID:9431187; PIDN:CAA24439.1; PID:915615  
A;Note: the authors did not translate the codon for residue 1  
C;Genetics:  
A;Gene: 18.7  
A;Map position: 92.73-93.35  
C;Superfamily: phage T7 gene 18.7 protein

Query Match 2.2%; Score 8; DB 1; Length 83;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 316 LSIKTLTP 323  
Db 22 LSIKTLTP 29

RESULT 6  
W8BP73  
gene 18.7 protein - phage T3  
C;Species: phage T3  
A;Note: host Escherichia coli  
C;Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 23-Jul-1999  
C;Accession: E23476  
R;Yamada, M.; Fujisawa, H.; Kato, H.; Hamada, K.; Minagawa, T.  
Virology 151, 350-361, 1986  
A;Title: Cloning and sequencing of the genetic right end of bacteriophage T3 DNA.  
A;Reference number: A94339; MUID:86209957  
A;Accession: E23476  
A;Molecule type: DNA  
A;Residues: 1-83 <YAM>  
A;Cross-references: GB:M14784; NID:9215810; PIDN:AA92527.1; PID:91196765  
R;Yamada, M.; Fujisawa, H.; Kato, H.; Hamada, K.; Minagawa, T.  
Virology 154, 246, 1986  
A;Reference number: A94344  
A;Contents: annotation; erratum; corrections to coding regions  
C;Genetics:  
A;Gene: 18.7  
C;Superfamily: phage T7 gene 18.7 protein

Query Match 2.2%; Score 8; DB 1; Length 83;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 316 LSIKTLTP 323  
Db 22 LSIKTLTP 29

RESULT 7  
H90892  
probable fimbrial chapone protein precursor [imported] - Escherichia coli (strain O  
C;Species: Escherichia coli

C.Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001  
 C.Accession: H90892  
 R.Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;  
 Gasawara, N.; Yatsunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.  
 DNA Res. 8, 11-22, 2001  
 A.Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and geno  
 A.Reference number: A99629; MUID:21156231; PMID:11258796  
 A.Accession: H90892  
 A.Status: preliminary  
 A.Molecule type: DNA  
 A.Residues: 1-236 <HAV>  
 A.Cross-references: GB:BA000007; PIDN:BA035535.1; PID:g13361578; GSPDB:GN00154  
 A.Experimental source: strain O157:H7, substrain RMD 0509952  
 C.Genetics:  
 A.Gene: ECs2112  
 C.Superfamily: chaperone protein papp

Query Match 2.2%; Score 8; DB 2; Length 236;  
 Best Local Similarity 100.0%; Pred. No. 4.2;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 204 LTPRIKAT 211  
 DB 227 LTPRIKAT 234

RESULT 8  
 B87524  
 1-acyl-sn-glycerol-3-phosphate acyltransferase [imported] - *Caulobacter crescentus*  
 C.Species: *Caulobacter crescentus*  
 C.Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
 C.Accession: B87524  
 R.Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
 B.; Land, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
 n, J.; Esmolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
 A.Title: Complete genome sequence of *Caulobacter crescentus*.  
 A.Reference number: A87249; MUID:21173698; PMID:11259647  
 A.Accession: B87524  
 A.Status: preliminary  
 A.Molecule type: DNA  
 A.Residues: 1-236 <STO>  
 A.Cross-references: GB:AE005673; NID:g13423724; PIDN:AAK24190.1; GSPDB:GN00148  
 C.Genetics:  
 A.Gene: CC2219

Query Match 2.2%; Score 8; DB 2; Length 236;  
 Best Local Similarity 100.0%; Pred. No. 4.2;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 169 VIFPECTR 176  
 DB 146 VIFPECTR 153

RESULT 9  
 A85725  
 probable fimbrial chaperone protein z2201 [imported] - *Escherichia coli* (strain O157:H7,  
 C.Species: *Escherichia coli*  
 C.Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
 C.Accession: A85725  
 R.Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
 Miller, L.; Grothbeck, E.J.; Davis, N.W.; Lam, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
 Nature 409, 529-533, 2001  
 A.Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.  
 A.Reference number: A85480; MUID:21074935; PMID:11206551  
 A.Accession: A85725  
 A.Status: preliminary  
 A.Molecule type: DNA  
 A.Residues: 1-239 <STO>  
 A.Cross-references: GB:AE005174; NID:g12515164; PIDN:AA656261.1; GSPDB:GN00145; UMGF:222

A.Experimental source: strain O157:H7, substrain EDL933  
 C.Genetics:  
 A.Gene: Z2201  
 C.Superfamily: chaperone protein papp

Query Match 2.2%; Score 8; DB 2; Length 239;  
 Best Local Similarity 100.0%; Pred. No. 4.3;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 204 LTPRIKAT 211  
 DB 230 LTPRIKAT 237

RESULT 10  
 A71478  
 probable metal dependent hydrolase - *Chlamydia trachomatis* (serotype D, strain UW3/Cx  
 C.Species: *Chlamydia trachomatis*  
 C.Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 08-Oct-1999  
 C.Accession: A71478  
 R.Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitche  
 Science 282, 754-759, 1998  
 A.Title: Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia t*  
 A.Reference number: A71570; MUID:99000809  
 A.Accession: A71478  
 A.Status: preliminary  
 A.Molecule type: DNA  
 A.Residues: 1-262 <ARN>  
 A.Cross-references: GB:AE001344; GB:AE001273; NID:g3329188; PIDN:AA68333.1; PID:g332  
 C.Genetics:  
 A.Experimental source: serotype D, strain UW-3/Cx  
 A.Gene: yycJ

Query Match 2.2%; Score 8; DB 2; Length 262;  
 Best Local Similarity 100.0%; Pred. No. 4.6;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 202 HVLTPRIK 209  
 DB 202 HVLTPRIK 209

RESULT 11  
 G71334  
 probable lysophosphatidic acid acyltransferase - *Syphilis spirochete*  
 C.Species: *Treponema pallidum* subsp. *pallidum* (*Syphilis spirochete*)  
 C.Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 05-Nov-1999  
 C.Accession: G71334  
 R.Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; G  
 rson, J.; Khalak, H.; Richardson, D.; Howell, J.R.; Chidambaram, M.; Uitterback, T.; M  
 they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.  
 Science 281, 375-388, 1998  
 A.Title: Complete genome sequence of *Treponema pallidum*, the *Syphilis spirochete*.  
 A.Reference number: A71250; MUID:98332770  
 A.Accession: G71334  
 A.Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A.Molecule type: DNA  
 A.Residues: 1-282 <COL>  
 A.Cross-references: GB:AE001215; GB:AE000520; NID:g3322631; PIDN:AA65346.1; PID:g332  
 A.Experimental source: strain Nichols  
 C.Genetics:  
 A.Gene: TP0361

Query Match 2.2%; Score 8; DB 2; Length 282;  
 Best Local Similarity 100.0%; Pred. No. 5;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 169 VIFPECTR 176  
 DB 170 VIFPECTR 177

## RESULT 12

S54641

probable membrane protein YDR018c - yeast (*Saccharomyces cerevisiae*)

N:Alternate names: hypothetical protein D3246; hypothetical protein P2F396; hypothetical

C:Species: *Saccharomyces cerevisiae*

C:Date: 08-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 20-Jun-2000

C:Accession: S54641; S63425; S67831; S72116

R:Pedman, K.; Brown, D.; Hamlyn, N.; Bowman, S.

Submitted to the EMBL Data Library, May 1995

A:Reference number: S54638

A:Accession: S54641

A:Molecule type: DNA

A:Residues: 1-396 &lt;DED&gt;

A:Cross-references: EMBL:Z49770; NID:G840867; PIDN:CAA9843.1; PID:G840871

A:Experimental source: strain AB972

R:Eide, L.G.; Sander, C.; Prydz, H.

Submitted to the EMBL Data Library, February 1996

A:Description: Sequencing and analysis of a 35.4 kb region on the left arm of chromosome

A:Reference number: S63416

A:Accession: S63425

A:Molecule type: DNA

A:Residues: 1-396 &lt;EID&gt;

A:Cross-references: EMBL:X95966; NID:G1216215; PIDN:CAA65210.1; PID:G1216225

R:Prydz, H.; Eide, L.G.

Submitted to the Protein Sequence Database, July 1996

A:Reference number: S67822

A:Accession: S67831

A:Molecule type: DNA

A:Residues: 1-396 &lt;PRY&gt;

A:Cross-references: EMBL:Z74314; NID:G1431443; PIDN:CAA9838.1; PID:G1431444; MIPS:YDR01

A:Experimental source: strain S288C

R:Eide, L.G.; Sander, C.; Prydz, H.

Yeast 12, 1085-1090, 1996

A:Title: Sequencing and analysis of a 35.4 kb region on the left arm of chromosome IV fr

A:Reference number: S72107; MIMD:97051598

A:Accession: S72116

A:Molecule type: DNA

A:Status: nucleic acid sequence not shown; translation not shown

A:Residues: 1-396 &lt;ETW&gt;

A:Cross-references: EMBL:X95966; NID:G1216215; PIDN:CAA65210.1; PID:G1216225

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1996

C:Genetics:

A:Map position: 4R

A:Map position: 4R

C:Superfamily: probable membrane protein YBR042c

C:Keywords: transmembrane protein

F:27-43/Domain: transmembrane #status predicted &lt;TM1&gt;

F:69-85/Domain: transmembrane #status predicted &lt;TM2&gt;

F:376-392/Domain: transmembrane #status predicted &lt;TM3&gt;

## RESULT 13

C87586

metal ion efflux membrane fusion protein family [Imported] - *Caulobacter crescentus*C:Species: *Caulobacter crescentus*

C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 30-Jun-2001

C:Accession: C87586

R:Merlan, W.C.; Feldblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; Debroy, R.T.; Durkin, R.J.; Gwin, M.L.; Haft, D.H.; Kolod

n, J.; Ernolova, M.; White, O.; Salberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of *Caulobacter crescentus*.

A:Reference number: A87249; MIMD:21173698; PMID:11259647

A:Accession: C87586

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-411 &lt;STO&gt;

A:Cross-references: GB:AE005673; NID:G13424311; PIDN:AAK24687.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC2722

C:Superfamily: nickel-cobalt resistance determinant structural protein CnrB; lipo1/b

Query Match

Best Local Similarity

Matches

Query Match

Best Local Similarity

Matches

Query Match

Best Local Similarity

Matches

Query Match

Best Local Similarity

Matches

Query Match

Best Local Similarity

Matches

Query Match

Best Local Similarity

Matches

Query Match

Best Local Similarity

Matches

Query Match

Best Local Similarity

Matches

Query Match

Best Local Similarity

Matches

Query Match

Best Local Similarity

Matches

Query Match

Best Local Similarity

Matches

Query Match

Best Local Similarity

Matches

Query Match

Best Local Similarity

Matches

Query Match

Best Local Similarity

Matches

Query Match

Best Local Similarity

Matches

Query Match

Best Local Similarity

Matches

Query Match

Best Local Similarity

Oy	178	NPEQTKVL	185
Db	291	NPEQTKVL	298

Search completed: August 28, 2002, 11:20:09  
Job time: 262 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:28:34 ; Search time 67.07 seconds

(without alignments)  
586.085 Million cell updates/sec

Title: US-09-853-526-70

Perfect score: 228

Sequence: 1 MRYLPSVVLGTAPTYVLA.....NYLDATYDVVVEGRKDDG 228

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 562222 seqs, 172994929 residues

Word size: 0

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

SPTREMBL\_19:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_unclassified:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query	Length	DB ID	Description
1	228	100.0	364	4	Q9BQ64
2	144	63.2	353	4	Q9N002
3	32	14.0	354	11	Q9DIE8
4	9	3.9	264	16	Q9PQ67
5	9	3.9	386	5	Q9SR12
6	8	3.5	115	16	Q9SE87
7	8	3.5	236	16	Q9A675
8	8	3.5	262	16	Q84743
9	8	3.5	282	16	Q83380
10	8	3.5	454	3	Q9P5W2
11	8	3.5	458	16	Q9KFI0
12	8	3.5	459	16	Q31503
13	8	3.5	483	10	Q49927
14	8	3.5	1820	5	Q9VVG0
15	7	3.1	54	6	Q28219
16	7	3.1	126	11	Q62370

17	7	3.1	132	11	Q9D581	Q9d581 mus musculus
18	7	3.1	137	12	Q55581	Q55581 leucania se
19	7	3.1	141	6	Q28218	Q28218 cercopithec
20	7	3.1	143	16	Q98R89	Q98R89 mycoplasma
21	7	3.1	148	16	Q9CH79	Q9ch79 lactococcus
22	7	3.1	150	2	Q9AGV0	Q9agv0 bruceella ab
23	7	3.1	162	10	Q9LPP0	Q9lpp0 arabidopsis
24	7	3.1	163	2	Q9EX02	Q9ex02 streptomyces
25	7	3.1	175	2	Q9EX12	Q9ex12 klebsiella
26	7	3.1	195	16	Q9KCD7	Q9kcd7 bacillus ha
27	7	3.1	209	16	Q91007	Q910u7 pseudomonas
28	7	3.1	209	16	Q9CJ4	Q9cjk4 pasteurella
29	7	3.1	211	16	Q67841	Q67841 aquifex aeo
30	7	3.1	220	2	Q9EMW7	Q9emw7 streptomyces
31	7	3.1	225	16	Q97NE9	Q97ne9 streptococcus
32	7	3.1	230	16	Q9PH25	Q9ph25 campylobact
33	7	3.1	234	2	Q32330	Q32330 clostridium
34	7	3.1	240	12	Q9WH72	Q9wh72 thailand to
35	7	3.1	240	16	P73054	P73054 synechocyst
36	7	3.1	241	16	Q97KF4	Q97kf4 clostridium
37	7	3.1	247	16	Q9XDI6	Q9xd16 neisseria m
38	7	3.1	247	16	Q9XDI4	Q9xd14 neisseria m
39	7	3.1	251	16	Q9KOA1	Q9k0a1 neisseria m
40	7	3.1	255	16	Q92LR6	Q92lr6 rhizobium m
41	7	3.1	257	16	Q917C1	Q917c1 pseudomonas
42	7	3.1	259	2	Q9F728	Q9f728 pseudomonas
43	7	3.1	262	16	Q9PLR8	Q9plr8 chlamydia m
44	7	3.1	270	5	Q9VDX1	Q9vdx1 dirosophila
45	7	3.1	272	16	Q98PM7	Q98pm7 mycoplasma

## ALIGNMENTS

RESULT	ID	PRELIMINARY:	PRT:	364 AA.
1	Q9BQ64			
Q9BQ64	Q9BQ64:			
AC	01-JUN-2001 (TREMBLrel. 17, Created)			
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	HYPOTHETICAL 42.1 KDA PROTEIN (LYSOPHOSPHATIDIC ACID			
DE	ACYLTRANSFERASE-EPSILON) (EC 2.3.1.51).			
GN	DKEP2761C222 OR LPAAT-E.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=AMYGDALA;			
RX	MEDLINE=21154917; PubMed=11230166;			
RA	Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,			
RA	Ansoerge W., Boecker M., Bloecher H., Bauersachs S., Blum H.,			
RA	Lauber J., Duesterhoeft A., Beyer A., Koehler K., Strack N.,			
RA	Mewes H.W., Oltjenwaelder B., Obermaier B., Tampe J., Heubner D.,			
RA	Wambutt R., Korn B., Klein M., Poustka A.;			
RT	"Towards a Catalog of Human Genes and Proteins: Sequencing and			
RT	Analysis of 500 Novel Complete Protein Coding Human cDNAs.";			
RT	Genome Res. 11:422-435(2001).			
[2]	SEQUENCE FROM N.A.			
RA	leung D.W.;			
RT	"Cloning and expression of LPAAT-epsilon.";			
RT	Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.			
DR	EMBL; AL136587; CAB65522.1; -			
DR	EMBL; AF375789; AAK54809.1; -			
DR	InterPro: IPR002123; Acyltransferase.			
DR	Pfam: PF01553; Acyltransferase; 1.			
KW	Hypothetical protein; transferase; Acyltransferase.			
SO	SEQUENCE 364 AA; 42072 MW; 90A0F87FC/78081 CRC64;			

```

Query Match      100.0%; Score 228; DB 4; Length 364;
Best Local Similarity 100.0%; Pred. No. 2.8e-234;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRLLPSVVLGTAPTYVLAMGWRLSAPLPRFYQALDDRLCYQSVWLEFFENYTG 60
   |||||||
DB 12 MRLLPSVVLGTAPTYVLAMGWRLSAPLPRFYQALDDRLCYQSVWLEFFENYTG 71
   |||||||

QY 61 VOILLVGDLPKKNENITYLANHOSYVDWIVADILAIQNALGHVRYLKEGLKWLPLYGC 120
   |||||||
DB 72 VOILLVGDLPKKNENITYLANHOSYVDWIVADILAIQNALGHVRYLKEGLKWLPLYGC 131
   |||||||

QY 121 YFAOHGIIYVRSRKNKLOSVDAGTPMYIVITEPSTRYNPEQTKVLSAQAF 180
   |||||||
DB 132 YFAOHGIIYVRSRKNKLOSVDAGTPMYIVITEPSTRYNPEQTKVLSAQAF 191
   |||||||

QY 181 AAOGLAVLKHVLTPTKATHVAFDCKKNYLDIVDTVYVEGKDDG 228
   |||||||
DB 192 AAOGLAVLKHVLTPTKATHVAFDCKKNYLDIVDTVYVEGKDDG 239
   |||||||

RESULT 2
Q9NUQ2 PRELIMINARY: PRT: 353 AA.
AC Q9NUQ2:
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE cDNA FLJ11210 F1S, CLONE 3LACE1007954.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCB1_Taxid=9606;
RN [1]
RP TISSUE=PLACENTA;
RC SEQUENCE FROM N.A.
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Wagatsuma M., Hosoliri T., Saku Y., Kodaira H., Kondo H., Sugawara M.,
RA Wakabayashi M., Chida Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Tanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuno Y.,
RA Minomiyagi K., Iwayanagi T.;
RT "NEO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK002072; BAA92069.1;
DR InterPro: IPR002123; Acyltransferase.
DR Pfam: PF01553; Acyltransferase; 1
SQ SEQUENCE 353 AA; 40799 MW; 0A8A1A42361F1B14 CRC64;

Query Match      63.2%; Score 144; DB 4; Length 353;
Best Local Similarity 100.0%; Pred. No. 8.6e-145;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRLLPSVVLGTAPTYVLAMGWRLSAPLPRFYQALDDRLCYQSVWLEFFENYTG 60
   |||||||
DB 1 MRLLPSVVLGTAPTYVLAMGWRLSAPLPRFYQALDDRLCYQSVWLEFFENYTG 60
   |||||||

QY 61 VOILLVGDLPKKNENITYLANHOSYVDWIVADILAIQNALGHVRYLKEGLKWLPLYGC 120
   |||||||
DB 61 VOILLVGDLPKKNENITYLANHOSYVDWIVADILAIQNALGHVRYLKEGLKWLPLYGC 120
   |||||||

QY 121 YFAOHGIIYVRSRKNKLOSVDAGTPMYIVITEPSTRYNPEQTKVLSAQAF 144
   |||||||
DB 121 YFAOHGIIYVRSRKNKLOSVDAGTPMYIVITEPSTRYNPEQTKVLSAQAF 144
   |||||||

RESULT 3
Q9DIE8 PRELIMINARY: PRT: 354 AA.
AC Q9DIE8:
DT 01-JUN-2001 (Tremblrel. 17, Created)

```

```

DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE 1110013A05RIK PROTEIN.
GN 1110013A05RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCB1_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
RX MEDLINE=21085660; Pubmed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Glass C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staudli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamija M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitlaker C., Wilming L.,
RA Wysshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK003649; BAB22915.1;
DR MGD: MGI:1915880; 1110013A05RIK.
DR InterPro: IPR002123; Acyltransferase.
DR Pfam: PF01553; Acyltransferase; 1
SQ SEQUENCE 354 AA; 40943 MW; 447EEF924B91E800 CRC64;

Query Match      14.0%; Score 32; DB 11; Length 354;
Best Local Similarity 100.0%; Pred. No. 1.9e-25;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 174 LSASQAFMAOGLAVLKHVLTPTKATHVAFD 205
   |||||||
DB 174 LSASQAFMAOGLAVLKHVLTPTKATHVAFD 205
   |||||||

RESULT 4
Q9POE7 PRELIMINARY: PRT: 264 AA.
AC Q9POE7:
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE.
GN PLSC OR UU344.
OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Ureaplasma.
OX NCB1_Taxid=134821;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SEROVAR 3;
RX MEDLINE=20500219; Pubmed=11048724;
RA Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,
RA Cassell G.H.;
RT "The complete sequence of the mucosal pathogen Ureaplasma
RT urealyticum.";
RL Nature 407:757-762(2000).
DR EMBL: AE002131; AAF30753.1;
DR InterPro: IPR002123; Acyltransferase.
DR Pfam: PF01553; Acyltransferase; 1.

```



KW Complete proteome.  
SQ SEQUENCE 264 AA; 30687 MW; DBC5727A07C74F7 CRC64;

Query Match  
Best Local Similarity 100.0%; Score 9; DB 16; Length 264;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 157 LVIPPEGR 165  
|||  
DB 156 LVIPPEGR 164

RESULT 5  
ID 095R12 PRELIMINARY; PRT; 386 AA.

AC 095R12; PRELIMINARY; PRT; 386 AA.  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE HYPOTHETICAL 44.0 KDA PROTEIN.  
GN F28B3.9.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_Taxid=6239;

RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RX MEDLINE=99069613; PubMed=9851916;  
RA None;  
RT "Genome sequence of the nematode C. elegans: a platform for  
investigating biology. The C. elegans Sequencing Consortium.";  
RL Science 282:2012-2018(1998).

RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA Geisel C., Kramer J., Smith A.;  
RT "The sequence of C. elegans cosmid F28B3.";  
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.

RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA Waterston R.;  
RT "Direct Submission.";  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF003136; AAK93853.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 386 AA; 43978 MW; A1E3020C201452A6 CRC64;

Query Match  
Best Local Similarity 100.0%; Score 9; DB 5; Length 386;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 211 LDATYDTV 219  
|||  
DB 230 LDATYDTV 238

RESULT 6  
ID 098E87 PRELIMINARY; PRT; 115 AA.

AC 098E87; PRELIMINARY; PRT; 115 AA.  
DT 01-OCT-2001 (TREMBlrel. 18, Created)  
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)  
DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)  
DE M14358 PROTEIN.  
GN M14358.

OS Rhizobium loti (Mesorhizobium loti).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Phyllobacteriaceae; Mesorhizobium.  
OX NCBI\_Taxid=381;

RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-MAFF303099;  
RX MEDLINE=21082930; PubMed=11214968;  
RA Kaneo T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,  
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,  
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,  
RA Takeuchi C., Yamada M., Tabata S.;  
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
Mesorhizobium loti.";  
RL DNA Res. 7:331-338(2000).  
DR EMBL; AP003004; BAB51033.1; -.  
KW Complete proteome.  
SQ SEQUENCE 115 AA; 12335 MW; 6AF43D6451F42BE8 CRC64;

Query Match  
Best Local Similarity 100.0%; Score 8; DB 16; Length 115;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 174 LSASOAF 181  
|||  
DB 31 LSASOAF 38

RESULT 7  
ID 09A675 PRELIMINARY; PRT; 236 AA.

AC 09A675; PRELIMINARY; PRT; 236 AA.  
DT 01-JUN-2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE.  
GN CC2219.  
OS Caulobacter crescentus.  
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
OC Caulobacter.  
OX NCBI\_Taxid=69394;

RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 19089 / CB15;  
RX MEDLINE=21173698; PubMed=11259647;  
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
RA Potocka L., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,  
RA Uterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,  
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
RT "Complete genome sequence of Caulobacter crescentus.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
DR EMBL; AE005893; AAK24190.1; -.  
DR TIGR; CC2219; -.  
DR InterPro; IPR002123; Acyltransferase.  
DR Pfam; PF01553; Acyltransferase; 1.  
KW Transferase; Acyltransferase; Complete proteome.  
SQ SEQUENCE 236 AA; 26362 MW; 85694D3966BCA1A0 CRC64;

Query Match  
Best Local Similarity 100.0%; Score 8; DB 16; Length 236;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 158 VIFPEGR 165  
|||  
DB 146 VIFPEGR 153

RESULT 8  
ID 084743 PRELIMINARY; PRT; 262 AA.

DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 09, Last annotation update)  
 DE METAL DEPENDENT HYDROLASI:  
 GN YECJ OR CT738.  
 OS Chlamydia trachomatis.  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
 NCBI\_TaxID=813;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=D/UV-3/CX;  
 RA MEDLINE=99000809; PubMed=9784136;  
 RA Stephens R.S., Kaiman S., Lammel C.J., Fan J., Marathe R., Aravind L.,  
 RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,  
 RA Davis R.W.;  
 RT "Genome sequence of an obligate intracellular pathogen of humans:  
 RT Chlamydia trachomatis."  
 RL Science 282:754-759(1998).  
 DR EMBL: AE001344; AAC68333.1; -.  
 KW Hydrolase; Complete proteome.  
 SO SEQUENCE 262 AA; 29082 MW; E36A6298A23CEFBF CRC64;

Query Match 3.5%; Score 8; DB 16; Length 262;  
 Best Local Similarity 100.0%; Pred. No. 5.4;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 191 HVLPTRIK 198  
 DB 202 HVLPTRIK 209

RESULT 9  
 ID 083380 PRELIMINARY; PRT; 282 AA.  
 AC 083380;  
 DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE LYOPHOSPHATIDIC ACID ACYLTRANSFERASE, PUTATIVE.  
 GN TP0361.  
 OS Treponema pallidum.  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.  
 NCBI\_TaxID=160;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NICHOLS;  
 RX MEDLINE=98332770; PubMed=9665876;  
 RX Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,  
 RX Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,  
 RX Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,  
 RX Khalak H., Richardson D., Howell J.K., Chidambaram M., Uterback T.,  
 RX McDonald L., Artlich P., Bowman C., Cotton M.D., Fujii C., Garland S.,  
 RX Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,  
 RX Venter J.C.;  
 RT "Complete genome sequence of Treponema pallidum, the syphilis  
 RT spirochete."  
 RL Science 281:375-388(1998).  
 DR EMBL: AE001215; AAC65346.1; -.  
 DR TIGR: TP0361; -.  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KW Transferase; Acyltransferase; Complete proteome.  
 SO SEQUENCE 282 AA; 31700 MW; DEF69003CD6C6AD CRC64;

Query Match 3.5%; Score 8; DB 16; Length 282;  
 Best Local Similarity 100.0%; Pred. No. 5.8;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 158 VIFPECTR 165  
 DB 170 VIFPECTR 177

RESULT 10  
 ID 09P5W2 PRELIMINARY; PRT; 454 AA.  
 AC 09P5W2;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE RELATED TO GUANINE DEAMINASE.  
 GN B3E4.200.  
 OS Neurospora crassa.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariales; Sordariaceae; Neurospora.  
 NCBI\_TaxID=5141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Schulte U., Algn V., Hohnsels J., Brandt P., Fartmann B., Holland R.,  
 RA Nyakatura G., Mewes H.W., Mannhaupt G.;  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA German Neurospora genome project;  
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL355931; CAB91408.2; -.  
 SO SEQUENCE 454 AA; 49778 MW; 8C1EC14BAECA421D CRC64;

Query Match 3.5%; Score 8; DB 3; Length 454;  
 Best Local Similarity 100.0%; Pred. No. 8.9;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 189 LKHVTPR 196  
 DB 208 LKHVTPR 215

RESULT 11  
 ID 09KF10 PRELIMINARY; PRT; 458 AA.  
 AC 09KF10;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE RNA METHYLTRANSFERASE.  
 GN BH0687.  
 OS Bacillus halodurans.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Bacillus.  
 NCBI\_TaxID=86665;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C-125 / JCM 9153;  
 RX MEDLINE=20512582; PubMed=11058132;  
 RX Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
 RX Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,  
 RX Horikoshi K.;  
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus  
 RT halodurans and genomic sequence comparison with Bacillus subtilis."  
 RL Nucleic Acids Res. 28:4317-4331(2000).  
 DR EMBL: AF001509; BAB04406.1; -.  
 DR InterPro: IPR000051; SAM\_bind.  
 DR InterPro: IPR001566; TRNA\_1.  
 DR PROSITE: PS01230; TRNA\_1; UNKNOWN\_1.  
 KW Transferase; Methyltransferase; Complete proteome.  
 SO SEQUENCE 458 AA; 51555 MW; 3D33A661B8891240 CRC64;

Query Match 3.5%; Score 8; DB 16; Length 458;  
 Best Local Similarity 100.0%; Pred. No. 8.9;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 167 NPEQTKVL 174

Db 291 NPEQTKVL 298

```

|||||||
RESULT 12
ID 031503 PRELIMINARY; PRT; 459 AA.
AC 031503;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, last sequence update)
DE 01-DEC-2001 (TREMBLrel. 18, last annotation update)
DE YEPF PROTEIN.
GN YEPF.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=96044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolochin A., Borchert S.,
RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Broillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Ertlan K.D., Erington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaere-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauder J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Meliado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogilawa A., Oudega B., Park S.H.,
RA Paro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowski E., Seror S.J., Serro P., Shi B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tanakoshi A., Tanaka T., Teapstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassaretti A.,
RA Viari A., Wanduit R., Wedler E., Wedler H., Weltzenegger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the gram-positive bacterium Bacillus
RT subtilis."
RL Nature 390:249-256(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
RL Submitted (NOV-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL: Z99107; CAB12493.1; -.
DR InterPro: IPR002792; DUF90.
DR InterPro: IPR000051; SAM_bind.
DR InterPro: IPR001566; TRMA_1.
DR Pfam: PF01938; TRAM; 1.
DR PROSITE: PS01230; TRMA_1; UNKNOWN_1.
DR PROSITE: PS01231; TRMA_2; 1.
KW Complete proteome.
SQ SEQUENCE 459 AA; 51819 MW; 35E9954779421D5 CRC64;

```

Query Match 3.5%; Score 8; DB 16; Length 459;  
 Best Local Similarity 100.0%; Pred. No. 9;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 167 NPEQTKVL 174

Db 292 NPEQTKVL 299

```

|||||||
RESULT 13
ID 049927 PRELIMINARY; PRT; 483 AA.
AC 049927;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE P54 PROTEIN.
GN P54.
OS Pisum sativum (garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX Submitted (FEB-1997) to the EMBL/Genbank/DBJ databases.
RA Castillo J., Marquez J.A., Franco L., Ballestar E., Rodrigo M.I.;
RL EMBL: Y11207; CAA72090.1; -.
DR HSSP: P50477; ICAU.
DR InterPro: IPR001113; Seedstore_7s.
DR Pfam: PF00546; Seedstore_7s; 1.
DR Pfam: PF02808; Seedstore_7s.C; 1.
SQ SEQUENCE 483 AA; 54662 MW; 8127BDAA0178F3D CRC64;

```

Query Match 3.5%; Score 8; DB 10; Length 483;  
 Best Local Similarity 100.0%; Pred. No. 9.4;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 ACTPMYLV 158  
 Db 184 ACTPMYLV 191

```

RESULT 14
ID Q9VVG0 PRELIMINARY; PRT; 1820 AA.
AC Q9VVG0;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)
DE PUTATIVE CADHERIN PRECURSOR (CG6445 PROTEIN).
GN CG6445.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinkner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanalides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Cheln L.X.,
RA Brandon R.C., Rogers Y.H.C., Blazej R.G., Champe C.M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson S.R., Miklos G.L.G.,
RA Abail J.F., Agdayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,

```

RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jamal M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Matel B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Mishina N.V., Modyarty C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacheb J.M.,  
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svitkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Maassman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Morley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*.";  
RL Science 287:2185-2195(2000).  
CC -1- SIMILARITY: BELONGS TO THE CADHERIN FAMILY.  
CC EMBL: AE003524; AAF49351.1; -;  
DR HSSP: P15116; INCU.  
DR FLYBase: FBgn0036715; CG6445.  
DR InterPro: IPR002126; Cadherin.  
DR Pfam: PF00028; cadherin.12.  
DR PRINTS: PR00205; CADHERIN.  
DR SMART: SM00112; CA1.13.  
DR PROSITE: PS00232; CADHERIN\_1; 1.  
DR PROSITE: PS50268; CADHERIN\_2; 13.  
KW Hypothetical protein; Cell adhesion; Glycoprotein; Repeat; Signal;  
KW Calcium-binding.  
FT SIGNAL 1 14  
FT CHAIN 15 1820 POTENTIAL.  
FT DOMAIN 36 128 POTATIVE CADHERIN.  
FT DOMAIN 137 245 CADHERIN 1.  
FT DOMAIN 254 365 CADHERIN 2.  
FT DOMAIN 374 483 CADHERIN 3.  
FT DOMAIN 492 585 CADHERIN 4.  
FT DOMAIN 590 696 CADHERIN 5.  
FT DOMAIN 718 815 CADHERIN 6.  
FT DOMAIN 824 939 CADHERIN 7.  
FT DOMAIN 955 1053 CADHERIN 8.  
FT DOMAIN 1062 1170 CADHERIN 9.  
FT DOMAIN 1170 1284 CADHERIN 10.  
FT DOMAIN 1295 1394 CADHERIN 11.  
FT DOMAIN 1407 1517 CADHERIN 12.  
FT DOMAIN 1517 1517 CADHERIN 13.  
FT CARBOHYD 191 191 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 371 371 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 435 435 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 678 678 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 727 727 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 868 868 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 890 890 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1040 1040 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1111 1111 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1112 1112 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1233 1233 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1294 1294 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1419 1419 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1452 1452 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1557 1557 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1584 1584 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1610 1610 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 1820 AA; 200996 MW; 8D61B907FA15C155 CRC64;

QY 102 GHVRYVLK 109  
Db 516 GHVRYVLK 523  
RESULT 15  
ID 028219 PRELIMINARY; PRT; 54 AA.  
AC 028219;  
DT 01-NOV-1996 (TEMBLrel. 01, Created)  
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)  
DE HEPARIN BINDING EGF-LIKE GENE PRECURSOR.  
OS Cercopithecus aethiops (Green monkey) (Grivet).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecinae; Cercopithecus.  
OX NCBI\_Taxid=9534;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=KIDNEY;  
RX MEDLINE=97444292; PubMed=9300824;  
RT Loukianov E.V., Loukianov T.I., Mledlocha A., Olunes S.;  
RT "Expression of mRNA for a short form of heparin-binding EGF-like  
RT growth factor.";  
RL Gene 195:81-86(1997).  
DR EMBL: X89729; CAA61881.1; -;  
KW Signal.  
FT SIGNAL 1 23 POTENTIAL.  
FT SEQUENCE 54 AA; 5888 MW; 7C0E67E85695873F CRC64;

Query Match 3.1%; Score 7; DB 6; Length 54;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LPSVVL 10  
Db 3 LPSVVL 9

Search completed: August 28, 2002, 11:28:35  
Job time: 548 sec

Query Match 3.5%; Score 8; DB 5; Length 1820;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

---

THIS PAGE BLANK (085470)

6)

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:20:09 ; Search time 39.16 Seconds  
(without alignments)  
559.458 Million cell updates/sec

Title: US-09-853-526-70

Perfect score: 228  
Sequence: 1 MRVLPSTVVLGTAPTYVLA.....NYLDAIVDVVVEGKDDGG 228

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 283138 seqs, 96089334 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

1: PIR\_71:\*  
2: PIR1:\*  
3: PIR2:\*  
4: PIR3:\*  
5: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	3.9	237	2	B71827
2	9	3.9	240	2	D64688
3	9	3.9	264	2	E82903
4	9	3.9	918	2	T34057
5	8	3.5	236	2	H90892
6	8	3.5	236	2	B87524
7	8	3.5	239	2	B85725
8	8	3.5	262	2	A71478
9	8	3.5	282	2	G71334
10	8	3.5	396	2	S54641
11	8	3.5	424	2	T49607
12	8	3.5	458	2	G83735
13	8	3.5	459	2	E69793
14	8	3.5	483	2	T06459
15	8	3.5	611	2	A85000
16	7	3.1	54	2	S58122
17	7	3.1	57	2	D86108
18	7	3.1	117	2	I81209
19	7	3.1	126	2	I81210
20	7	3.1	141	2	S58117
21	7	3.1	143	2	B90523
22	7	3.1	148	2	C86732
23	7	3.1	194	2	PC1136
24	7	3.1	195	2	C83854
25	7	3.1	208	1	A38432
26	7	3.1	208	1	A41914
27	7	3.1	208	1	UC1409
28	7	3.1	209	2	B83329
29	7	3.1	210	2	A10342

30	7	3.1	211	2	E70476	2-acylglycerophosph
31	7	3.1	225	2	H95244	conserved hypothet
32	7	3.1	226	2	E98109	conserved hypothet
33	7	3.1	230	2	E81397	probable 1-acylgly
34	7	3.1	240	2	S75162	hypothetical prote
35	7	3.1	241	2	B97019	1-acyl-sn-glycerol
36	7	3.1	247	2	A81957	1-acyl-glycerol-3-p
37	7	3.1	247	2	G81013	1-acyl-sn-glycerol
38	7	3.1	250	2	T27772	26S proteasome reg
39	7	3.1	250	2	E70104	1-acylglycerol-3-p
40	7	3.1	250	2	AF3384	outer membrane pro
41	7	3.1	251	2	A81167	conserved hypothet
42	7	3.1	257	2	A83645	probable acyltrans
43	7	3.1	262	2	H81740	conserved hypothet
44	7	3.1	268	2	AH2521	hypothetical prote
45	7	3.1	269	2	G91169	probable acyltrans

#### ALIGNMENTS

RESULT 1  
B71827  
Probable 1-acylglycerol-3-phosphate O-acyltransferase (EC 2.3.1.51) - Helicobacter pylori  
C:Species: Helicobacter pylori  
A:Variate: strain J99  
C:Date: 12-Feb-1999 #sequence-revision 12-Feb-1999 #text-change 06-Oct-2000  
C:Accession: B71827  
R:Alm, R.A.; Ling, L.S.L.; Molr, D.T.; King, B.L.; Brown, E.D.; Doly, P.C.; Smith, D.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.  
Nature 397, 176-180, 1999  
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric P.  
A:Reference number: A71800; M0ID:99120557  
A:Accession: B71827  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-237 <ARN>  
A:Cross-references: GB:AE001550; GB:AE001439; NID:g4155872; PIDN:AA06852.1; PID:g415  
A:Experimental source: strain J99  
C:Genetics:  
A:Gene: plsc  
C:Superfamily: mouse 1-acylglycerol-3-phosphate O-acyltransferase  
C:Keywords: acyltransferase; coenzyme A

Query Match 3.9%; Score 9; DB 2; Length 237;  
Best Local Similarity 100.0%; Pred. No. 0.27;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 LVTFPEGTR 165  
DB 142 LVTFPEGTR 150

RESULT 2  
D64688  
Probable 1-acylglycerol-3-phosphate O-acyltransferase (EC 2.3.1.51) - Helicobacter pylori  
C:Species: Helicobacter pylori  
C:Date: 09-Aug-1997 #sequence-revision 09-Aug-1997 #text-change 06-Oct-2000  
C:Accession: D64688  
R:Tomp, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.  
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKee  
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watney,  
Nature 388, 539-547, 1997  
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpp, P.D.; Smith, H.O.; Fraser,  
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A:Reference number: A64520; M0ID:97394467  
A:Accession: D64688  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-240 <TOM>  
A:Cross-references: GB:AE000636; GB:AE000511; NID:g2114517; PIDN:AA08393.1; PID:g231  
C:Superfamily: mouse 1-acylglycerol-3-phosphate O-acyltransferase

C;Keywords: acyltransferase; coenzyme A

Query Match 3.9%; Score 9; DB 2; Length 240;  
Best Local Similarity 100.0%; Pred. No. 0.27;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 157 LVIRPECTR 165  
|||||  
Db 142 LVIRPECTR 150

#### RESULT 3

1-acyl-sn-glycerol-3-phosphate acyltransferase UN344 [imported] - Ureaplasma urealyticum  
C:Species: Ureaplasma urealyticum  
C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C:Accession: E82903  
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Helner, C.R.; Chen, E.Y.; Cassell, G.H.  
Submitted to GenBank, February 2000  
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mlr  
A:Reference number: A82870  
A:Accession: E82903  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-264 <GUA>  
A:Cross-references: GB:AE002131; GB:AF222894; NID:96899316; PIDN:AAF30753.1; GSPDB:GN001  
A:Experimental source: serovar 3; biovar 1  
C:Genetics:  
A:Gene: pLSC: UN344  
A:Genetic code: SGC3

Query Match 3.9%; Score 9; DB 2; Length 264;  
Best Local Similarity 100.0%; Pred. No. 0.3;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 157 LVIRPECTR 165  
|||||  
Db 156 LVIRPECTR 164

#### RESULT 4

hypothetical protein F28B3.5 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
C:Accession: T34057  
R:Geisler, C.; Kramer, J.; Smith, A.  
Submitted to the EMBL Data Library, May 1997  
A:Description: The sequence of C. elegans cosmid F28B3.  
A:Reference number: 221469  
A:Accession: T34057  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-918 <GEI>  
A:Cross-references: EMBL:AF003136; PIDN:AA893636.1; GSPDB:GN00019; CESP:F28B3.5  
A:Experimental source: strain Bristol N2; clone F28B3  
C:Genetics:  
A:Gene: CESP:F28B3.5  
A:Map position: 1  
A:Intons: 85/3; 129/3; 235/3; 418/2; 482/3; 532/3; 736/3; 829/3; 850/2

Query Match 3.9%; Score 9; DB 2; Length 918;  
Best Local Similarity 100.0%; Pred. No. 0.9;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 211 LDAIVDVT 219  
|||||  
Db 762 LDAIVDVT 770

RESULT 5  
H90892  
Probable fimbrial chaperone protein precursor [imported] - Escherichia coli (strain O

C:Species: Escherichia coli  
C>Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001  
C:Accession: H90892  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shindagawa, H.  
DNA Res. 8, 11-22, 2001  
A>Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g  
A:Reference number: A99629; M01D:21156231; PMID:11258796  
A:Accession: H90892  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-236 <NAV>  
A:Cross-references: GB:BA000007; PIDN:BA835535.1; PID:q13361578; GSPDB:GN00154  
A:Experimental source: strain O157:H7, substrain RIMD 0509952  
C:Genetics:  
A:Gene: ECG2112  
C:Superfamily: chaperone protein papd

Query Match 3.5%; Score 8; DB 2; Length 234;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0;

OY 193 LTPRIKAT 200  
|||||  
Db 227 LTPRIKAT 234

#### RESULT 6

B87524  
1-acyl-sn-glycerol-3-phosphate acyltransferase [imported] - Caulobacter crescentus  
C:Species: Caulobacter crescentus  
C>Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
C:Accession: B87524  
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A>Title: Complete Genome Sequence of Caulobacter crescentus.  
A:Reference number: A87249; M01D:21173698; PMID:11259647  
A:Accession: B87524  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-236 <STO>  
A:Cross-references: GB:AE005673; NID:q13423724; PIDN:AAK24190.1; GSPDB:GN00148  
C:Genetics:  
A:Gene: CC2219

Query Match 3.5%; Score 8; DB 2; Length 236;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 158 VIFPECTR 165  
|||||  
Db 146 VIFPECTR 153

#### RESULT 7

A85725  
Probable fimbrial chaperone protein 22201 [imported] - Escherichia coli (strain O157:  
C:Species: Escherichia coli  
C>Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
C:Accession: A85725  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May iller, L.; Grobeck, E.J.; Davis, N.W.; Linn, A.; Dimalanta, E.; Potamousis, K.; Apoda Nature 409, 529-533, 2001  
A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480; M01D:21074935; PMID:11206551  
A:Accession: A85725

THIS PAGE IS BLANK (USP)



A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-239 <STO>  
A:Cross-references: GB:AE005174; NID:g12515164; PIDN:AAG56261.1; GSPDB:GN00145; UMGCP:222  
A:Experimental source: strain 0157:H7, substrain EDJ933  
C:Genetics:  
A:Gene: 22201  
C:Superfamily: chaperone protein papp

Query Match 3.5%; Score 8; DB 2; Length 239;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 193 LTPPRKAT 200  
DB 230 LTPPRKAT 237

#### RESULT 8

A71478

probable metal dependent hydrolase - Chlamydia trachomatis (serotype D, strain UW3/Cx)

C:Species: Chlamydia trachomatis

C:Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 08-Oct-1999

C:Accession: A71478

R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell,

Science 282, 754-759, 1998

A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trac

A:Reference number: A71570; MUID:99000809

A:Accession: A71478

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-262 <ARN>

A:Cross-references: GB:AE001344; GB:AE001273; NID:g3329188; PIDN:AAC68333.1; PID:g332919

A:Experimental source: serotype D, strain UW-3/Cx

C:Genetics:

A:Gene: yycJ

Query Match 3.5%; Score 8; DB 2; Length 262;  
Best Local Similarity 100.0%; Pred. No. 3.3;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 191 HVLTPRIK 198  
DB 202 HVLTPRIK 209

#### RESULT 9

G71334

probable lysophosphatidic acid acyltransferase - syphilis spirochete

C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)

C:Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 05-Nov-1999

C:Accession: G71334

R:Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin

ton, J.; Khailak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McD

they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.

Science 281, 375-388, 1998

A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.

A:Reference number: A71250; MUID:9832770

A:Accession: G71334

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-282 <COT>

A:Cross-references: GB:AE001215; GB:AE000520; NID:g3322631; PIDN:AAG55346.1; PID:g332263

A:Experimental source: strain Nichols

C:Genetics:

A:Gene: TP0361

OY 158 VIFPECTR 165  
DB 170 VIFPECTR 177

#### RESULT 10

S54641

probable membrane protein YDR018c - yeast (Saccharomyces cerevisiae)

M:Alternate names: hypothetical protein D3246; hypothetical protein PF396; hypotheti

C:Species: Saccharomyces cerevisiae

C:Date: 08-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 20-Jun-2000

C:Accession: S54641; S63425; S67831; S72116

R:Dedman, K.; Brown, D.; Hamlyn, N.; Bowman, S.

submitted to the EMBL Data Library, May 1995

A:Reference number: S54638

A:Accession: S54641

A:Molecule type: DNA

A:Residues: 1-396 <DED>

A:Cross-references: EMBL:Z49770; NID:g840867; PIDN:CAA69843.1; PID:g840871

A:Experimental source: strain AB972

R:Eide, L.G.; Sander, C.; Prydz, H.

submitted to the EMBL Data Library, February 1996

A:Description: Sequencing and analysis of a 35.4 kb region on the left arm of chromos

A:Reference number: S63416

A:Accession: S63425

A:Molecule type: DNA

A:Residues: 1-396 <EID>

A:Cross-references: EMBL:X95966; NID:g1216215; PIDN:CAA65210.1; PID:g1216225

R:Prydz, H.; Eide, L.G.

submitted to the Protein Sequence Database, July 1996

A:Reference number: S67822

A:Accession: S67831

A:Molecule type: DNA

A:Residues: 1-396 <PRX>

A:Cross-references: EMBL:Z74314; NID:g1431443; PIDN:CAA98838.1; PID:g1431444; MIPS:YD

A:Experimental source: strain S288C

R:Eide, L.G.; Sander, C.; Prydz, H.

Yeast 12, 1085-1090, 1996

A:Title: Sequencing and analysis of a 35.4 kb region on the left arm of chromosome IV

A:Reference number: S72107; MUID:97051598

A:Accession: S72116

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-396 <EIV>

A:Cross-references: EMBL:X95966; NID:g1216215; PIDN:CAA65210.1; PID:g1216225

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1996

C:Genetics:

A:Map position: 4R

A:Note: YDR018C

C:Superfamily: probable membrane protein YBR042C

C:Keywords: transmembrane protein

F:27-43/Domain: transmembrane #status predicted <TM1>

F:69-85/Domain: transmembrane #status predicted <TM2>

F:376-392/Domain: transmembrane #status predicted <TM3>

Query Match 3.5%; Score 8; DB 2; Length 396;  
Best Local Similarity 100.0%; Pred. No. 4.7;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 211 LDAIYDVT 218  
DB 265 LDAIYDVT 272

#### RESULT 11

T49607

guanine deaminase (Gda) related protein [Imported] - Neurospora crassa

M:Alternate names: protein B3E4.200

C:Species: Neurospora crassa

C:Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2000

C:Accession: T49607

R;Schulte, U.; Algn, V.; Hehnel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura, submitted to the Protein Sequence Database, May 2000  
A:Reference number: 225022  
A:Accession: T49607  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-424 <SCH>  
A:Cross-references: EMBL:AJ355931; GSPDB:GN00116; NCSP:B3E4.200  
A:Experimental source: BAC clone B3E4; strain OR74A  
C:Genetics:  
A:Gene: NCSP:B3E4.200  
A:Map position: 6

Query Match 3.5%; Score 8; DB 2; Length 424;  
Best Local Similarity 100.0%; Pred. No. 5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 189 LKHVLTFR 196  
Db 178 LKHVLTFR 185

RESULT 12  
G83735  
RNA methyltransferase BH0687 [imported] - Bacillus halodurans (strain C-125)  
C:Species: Bacillus halodurans  
C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
C:Accession: G83735  
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and A:Reference number: A83650; MID:20512582; PMID:11058132  
A:Accession: G83735  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-458 <STO>  
A:Cross-references: GB:AP001509; GB:BA000004; NID:g10173176; PIDN:BAB04406.1; GSPDB:GN00 A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: BH0687

Query Match 3.5%; Score 8; DB 2; Length 458;  
Best Local Similarity 100.0%; Pred. No. 5.3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 167 NPEOTKVL 174  
Db 291 NPEOTKVL 298

RESULT 13  
E69793  
RNA methyltransferase homolog yefa - Bacillus subtilis  
C:Species: Bacillus subtilis  
C:Date: 03-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 20-Jun-2000  
C:Accession: E69793  
R;Kunst, F.; Ogaewara, N.; Moszer, I.; Albertini, A.M.; Allioni, G.; Azevedo, V.; Bierter C.; Bron, S.; Brouillet, S.; Hruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd A.; Ehlich, S.D.; Emerson, P.T.; Ettlin, K.D.; Errington, J.; Fabrev, C.; Ferrati, E. Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F. Koetler, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle Rieger, M.; Rivoita, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serod akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpestra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K A:Authors: Yoshikawa, H.F.; Zumslein, E.; Yoshikawa, H.; Danchin, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A:Reference number: A69580; MID:98044033  
A:Accession: E69793  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-459 <KUN>  
A:Cross-references: GB:299107; GB:AL009126; NID:g2632866; PIDN:CAB12493.1; PID:g26329 A:Experimental source: strain 168  
C:Genetics:  
A:Gene: yefa  
C:Superfamily: hypothetical protein HI0333

Query Match 3.5%; Score 8; DB 2; Length 459;  
Best Local Similarity 100.0%; Pred. No. 5.3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 167 NPEOTKVL 174  
Db 292 NPEOTKVL 299

RESULT 14  
T06459  
62K sucrose-binding protein homolog - garden pea  
N:Alternate names: P54 protein  
C:Species: Pisum sativum (garden pea)  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 20-Jun-2000  
C:Accession: T06459  
R;Rodrigo II, M.I.  
submitted to the EMBL Data Library, February 1997  
A:Reference number: Z15693  
A:Accession: T06459  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-483 <ROD>  
A:Cross-references: EMBL:Y11207; PIDN:CA72090.1  
A:Experimental source: var. Lincoln  
C:Genetics:  
A:Note: P54  
C:Superfamily: glycinin

Query Match 3.5%; Score 8; DB 2; Length 483;  
Best Local Similarity 100.0%; Pred. No. 5.6;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 151 AGTPMYLV 158  
Db 184 AGTPMYLV 191

RESULT 15  
A85000  
heat shock protein hscA [imported] - Buchnera sp. (strain APS)  
C:Species: Buchnera sp.  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 23-Mar-2001  
C:Accession: A85000  
R;Chienobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H. Nature 407, 81-86, 2000  
A:Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp A:Reference number: A84930; MID:20445173  
A:Accession: A85000  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-611 <STO>  
A:Cross-references: GB:AP000398; GSPDB:GN00144  
A:Experimental source: strain APS  
C:Genetics:  
A:Gene: hscA; B0605  
C:Superfamily: heat shock protein 70

Query Match 3.5%; Score 8; DB 2; Length 611;

Best Local Similarity 100.0%; Pred. No. 6.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYLLPSV 9  
          |||||||  
Db 42 RYLLPSV 49

Search completed: August 28, 2002, 11:20:09  
Job time: 262 sec

**THIS PAGE BLANK (uspto)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:18:47 ; Search time 21.78 Seconds

(without alignments)  
627.548 Million cell updates/sec

Title: US-09-853-526-4

Perfect score: 353  
Sequence: 1 MRVLLPSVVLGTAPTYVLA.....YVNTWIGTLLGLMTVTKA 353

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size : 0

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	353	100.0	353	1 PICE_HUMAN	Q9nq22 homo sapien
2	32	9.1	354	1 PICE_MOUSE	Q9die8 mus musculu
3	9	2.5	237	1 PLSC_HELPI	Q923n8 helicobacte
4	9	2.5	240	1 PLSC_HELPI	Q25903 helicobacte
5	8	2.3	83	1 V187_BPT3	P10302 bacterioph
6	8	2.3	86	1 V187_BPT3	P03788 bacterioph
7	8	2.3	396	1 YD18_YEAST	Q12185 saccharomy
8	8	2.3	611	1 HSCA_BUCAL	P57660 buchnera ap
9	7	2.0	57	1 YJDO_ECOLI	P58638 escherichia
10	7	2.0	194	1 PURI_LACCA	P35853 lactobacill
11	7	2.0	208	1 HBGF_CERAE	Q09118 cercopithe
12	7	2.0	208	1 HBGF_HUMAN	Q99075 homo sapien
13	7	2.0	208	1 HBGF_PIG	Q01580 sus scrofa
14	7	2.0	208	1 HBGF_PIG	Q06175 ratus norv
15	7	2.0	250	1 HXB4_MOUSE	P10284 mus musculu
16	7	2.0	250	1 PLSC_BORBU	Q59188 borrelia bu
17	7	2.0	250	1 PSD8_CAEEL	Q23449 caenorhabd
18	7	2.0	251	1 HXB4_HUMAN	P17483 homo sapien
19	7	2.0	286	1 PROC_YEAST	P32265 saccharomy
20	7	2.0	308	1 PLSC_COCON	Q42670 cocos nucif
21	7	2.0	311	1 MPCP_YEAST	P23641 saccharomy
22	7	2.0	423	1 ENO_MERYA	Q60173 methanococ
23	7	2.0	445	1 XTLA_BACSU	P04780 bacillus su
24	7	2.0	516	1 GUX1_PCHAC	P13680 phanerocha
25	7	2.0	520	1 CMCH_NOCCLA	Q51080 nocardia la
26	7	2.0	532	1 GSI_NEUCR	P38678 neuropept
27	7	2.0	677	1 Y593_TREPA	Q83602 treponema p
28	7	2.0	860	1 CH12_COCHI	P54197 coccidioid
29	7	2.0	904	1 Y002_CAEEL	Q09228 caenorhabd
30	7	2.0	1018	1 ST31_MOUSE	Q992m1 mus musculu
31	7	2.0	1019	1 ST31_HUMAN	Q9bku1 homo sapien
32	7	2.0	1026	1 BGAL_STRTR	P23989 streptococ
33	7	2.0	1043	1 SYL_MERTH	Q27428 methanobact

34	7	2.0	1044	1 SYL_MERTH	P26499 methanobact
35	7	2.0	2485	1 PTND_HUMAN	Q12923 homo sapien
36	6	1.7	25	1 PCW2_PACGO	P82423 pachycondyl
37	6	1.7	25	1 PCW2_PACGO	P82423 pachycondyl
38	6	1.7	71	1 YAAA_BACSU	P05650 bacillus su
39	6	1.7	76	1 BB11_SCHCO	P78742 schizophyll
40	6	1.7	83	1 RL23_HALHA	Q06842 halobacteri
41	6	1.7	87	1 Y1WC_BPPH1	P10435 bacterioph
42	6	1.7	88	1 PMRD_ECOLI	P37590 escherichia
43	6	1.7	88	1 Y4OL_RHISN	P55597 rhizobium s
44	6	1.7	99	1 RPOL_METVA	Q57832 methanococ
45	6	1.7	101	1 THST_THETS	Q9nbw4 thermozymon

## ALIGNMENTS

RESULT	1	STANDARD	PRT	353 AA.
PLCE_HUMAN				
ID	PLCE_HUMAN			
AC	Q9N0G4			
DT	01-MAR-2002 (Rel. 41, Created)			
DT	01-MAR-2002 (Rel. 41, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	1-acyl-sn-glycerol-3-phosphate acyltransferase epsilon (EC 2.3.1.51)			
DE	(1-Agp acyltransferase 5) (1-AGPAT 5) (Lysophosphatidic acid			
DE	acyltransferase-epsilon) (LPAAT-epsilon) (1-acylglycerol-3-phosphate			
DE	O-acyltransferase 5).			
GN	AGPAT5.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Leung D.W.;			
RT	"Cloning and expression of LPAAT-epsilon.";			
RL	Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Cohen D., Chumakov I., Blumenfeld M., Bougueleret L.;			
RL	Patent number WO9332644, 01-JUL-1999.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	TISSUE-Amygdala;			
RX	MEDLINE=21154917; PubMed=11230166;			
RA	Wiemann S., Weill B., Wellenreuther R., Gassenhuber J., Glassl S.,			
RA	Ansorge W., Boecker M., Bloeker H., Bauersachs S., Blum H.,			
RA	Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,			
RA	Mewes H.-W., Othenwaelder B., Obermaier B., Tampe J., Heubner D.,			
RA	Wandutt R., Korn B., Klein M., Poustka A.;			
RT	"Towards a catalog of human genes and proteins: sequencing and			
RL	analysis of 500 novel complete protein coding human cDNAs.";			
RA	Genome Res. 11:422-435(2001).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RA	TISSUE=Placenta;			
RA	Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,			
RA	Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,			
RA	Wagatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,			
RA	Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,			
RA	Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saio K.,			
RA	Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y.,			
RT	"NEO human cDNA sequencing project.";			
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.			
CC	- FUNCTION: CONVERTS LYSOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC			
CC	ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION (BY			
CC	SIMILARITY).			
CC	- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate =			
CC	CoA + 1,2-diacyl-sn-glycerol 3-phosphate.			
CC	- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.			
CC	- SUBCELLULAR LOCATION: Integral membrane protein (potential).			

CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
 CC ACYLTRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF375789; AAK54809.1; ALT\_INT.  
 DR EMBL: AL136587; CAB66522.1; ALT\_INT.  
 DR EMBL: AK002072; BAA92069.1; -  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KM Phospholipid biosynthesis; Transferase; Acyltransferase;  
 KM Transmembrane.  
 FT TRANSMEM 7 29 POTENTIAL.  
 FT TRANSMEM 44 66 POTENTIAL.  
 FT TRANSMEM 334 351 POTENTIAL.  
 FT CONFLICT 145 145 L -> V (IN REF. 2).  
 SQ SEQUENCE 353 AA; 40813 MW; A05B1FA246CE1B64 CRC64;

Query Match 100.0%; Score 353; DB 1; Length 353;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVVLGTAPTYVLANGVRLLSAPLPAFYQALDRLCYQVQSWLFFFEWYTG 60  
 DB 1 MRYLLPSVVLGTAPTYVLANGVRLLSAPLPAFYQALDRLCYQVQSWLFFFEWYTG 60  
 QY 61 VOILLYGDLPEKNENIYLANHSTVDIYADIAIRQNALGHRYVLEKGLMPLPGC 120  
 DB 61 VOILLYGDLPEKNENIYLANHSTVDIYADIAIRQNALGHRYVLEKGLMPLPGC 120  
 QY 121 YFAOHGIIYKRSKAFNEKEKRNKLOSYVDAGTPMYLVEPEGRYNEPQTKVLSASQAF 180  
 DB 121 YFAOHGIIYKRSKAFNEKEKRNKLOSYVDAGTPMYLVEPEGRYNEPQTKVLSASQAF 180  
 QY 181 AAOAGGLAVLKHVLTPIKATPAVADCKKNYLDATYDVYVYEGKDDGQRESPTMTEFL 240  
 DB 181 AAOAGGLAVLKHVLTPIKATPAVADCKKNYLDATYDVYVYEGKDDGQRESPTMTEFL 240  
 QY 241 CKESPKIHIDRIKDVPEEOEHMRRLHERFEIKDKMLIEFYESPDPERRRRFGKS 300  
 DB 241 CKESPKIHIDRIKDVPEEOEHMRRLHERFEIKDKMLIEFYESPDPERRRRFGKS 300  
 QY 301 VNSKLSIKTKLPSMLILSGLTASMLMTDAGRKLYVNTWYIGTLGCLMWTIKA 353  
 DB 301 VNSKLSIKTKLPSMLILSGLTASMLMTDAGRKLYVNTWYIGTLGCLMWTIKA 353  
 RESULT 2  
 PLSC\_MOUSE STANDARD: PRT; 354 AA.  
 AC 09D1E8;  
 DT 01-MAR-2002 (Rel. 41, Created)  
 DT 01-MAR-2002 (Rel. 41, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE 1-acyl-sn-glycerol-3-phosphate acyltransferase epsilon (EC 2.3.1.51)  
 DE (1-AGP acyltransferase 5) (1-AGPAT 5) (Lysophosphatidic acid  
 DE acyltransferase-epsilon) (LPAT-epsilon) (1-acylglycerol-3-phosphate  
 DE O-acyltransferase 5).  
 OS AGPAT5.  
 GN Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Embryo;

RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shingawa A., Shibata Y., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Atzawa K., Iwaza M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Glasi C., King B., Kochiwa H.,  
 RA Kiehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
 RA Schirral L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Mashio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carrincci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gwinnich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seye T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Wetz C., Whitaker C., Wilming L.,  
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kontsuki S.,  
 RA Hayashizaki Y.,  
 RA "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 CC -1- FUNCTION: CONVERTS LYSOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC  
 CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION (BY  
 CC SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate =  
 CC CoA + 1,2-diacyl-sn-glycerol 3-phosphate.  
 CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
 CC ACYLTRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AK003649; BAB22915.1; -  
 DR MGD: MGI:1915880; Agpat5.  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KM Phospholipid biosynthesis; Transferase; Acyltransferase;  
 KM Transmembrane.  
 FT TRANSMEM 7 29 POTENTIAL.  
 FT TRANSMEM 44 66 POTENTIAL.  
 FT TRANSMEM 335 352 POTENTIAL.  
 SQ SEQUENCE 354 AA; 40943 MW; 447EEF924B91E800 CRC64;

Query Match 9.1%; Score 32; DB 1; Length 354;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-24;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 174 LSASQAFRAQRCGLAVLKHVLTPIKATHVAFD 205  
 DB 174 LSASQAFRAQRCGLAVLKHVLTPIKATHVAFD 205  
 RESULT 3  
 PLSC\_HELPJ STANDARD: PRT; 237 AA.  
 AC 09ZJN6;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE 1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51) (1-AGP  
 DE acyltransferase) (1-AGPAT) (Lysophosphatidic acid acyltransferase)  
 DE (LPATV).  
 OS PLSC OR JHP1267.  
 GN Helicobacter pylori j99 (Campylobacter pylori j99).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;

```

OC Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Camel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori."
RL Nature 357:176-180(1999).
CC -1- FUNCTION: CONVERTS LYSOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC
CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION.
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate =
CC CoA + 1,2-diacyl-sn-glycerol 3-phosphate.
CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).
CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE
CC ACYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE001550; AAD06852.1; -
DR InterPro: IPR002123; Acyltransferase.
DR Pfam: PF01553; Acyltransferase; 1.
KW Phospholipid biosynthesis; Transferase; Acyltransferase;
KW Inner membrane; Complete proteome.
SQ SEQUENCE 237 AA; 27188 MW; E10F517D42A1731F CRC64;

Query Match 2.5%; Score 9; DB 1; Length 237;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 LVTFPEGTR 165
DB 142 LVTFPEGTR 150

RESULT 4
PLSC_HELPY STANDARD; PRT; 240 AA.
AC 025903;
DT 15-DEC-1998 (rel. 37, Created)
DT 15-DEC-1998 (rel. 37, Last sequence update)
DT 16-OCT-2001 (rel. 40, Last annotation update)
DE 1-acyl-sn-glycerol 3-phosphate acyltransferase (EC 2.3.1.51) (1-AGP
DE acyltransferase) (1-AGPAT) (Lysophosphatidic acid acyltransferase)
DE (LPAAT).
GN PLSC OR HP1348.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kervilange A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McEnaney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Meldman J.M., Fujii C., Bowman C., Wathey L., Wallin E.,

```

```

RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori."
RL Nature 388:539-547(1997).
CC -1- FUNCTION: CONVERTS LYSOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC
CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION.
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate =
CC CoA + 1,2-diacyl-sn-glycerol 3-phosphate.
CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).
CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE
CC ACYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000636; AAD08393.1; -
DR TIGR: HP1348; -
DR InterPro: IPR002123; Acyltransferase.
DR Pfam: PF01553; Acyltransferase; 1.
KW Phospholipid biosynthesis; Transferase; Acyltransferase;
KW Inner membrane; Complete proteome.
SQ SEQUENCE 240 AA; 27745 MW; 22BD5D0EB190BBD CRC64;

Query Match 2.5%; Score 9; DB 1; Length 240;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 LVTFPEGTR 165
DB 142 LVTFPEGTR 150

RESULT 5
V187_BP73 STANDARD; PRT; 83 AA.
AC P10302;
DT 01-MAR-1989 (rel. 10, Created)
DT 01-MAR-1989 (rel. 10, Last sequence update)
DT 01-MAR-1989 (rel. 10, Last annotation update)
DE Gene 18.7 protein.
GN 18.7.
OS Bacteriophage T3.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC T7-like phages.
OX NCBI_TaxID=10759;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86209997; PubMed=3010556;
RA Yamada M., Fujisawa H., Kato H., Hamada K., Minagawa T.;
RT "Cloning and sequencing of the genetic right end of bacteriophage T3
RT DNA."
RL Virology 151:350-361(1986).
RN [2]
RP ERRATUM.
RA Yamada M., Fujisawa H., Kato H., Hamada K., Minagawa T.;
RL Virology 154:246-246(1986).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

DR EMBL: M14784; AAA92527.1; -  
 DR PIR: E23476; WBBP73.  
 SQ SEQUENCE 83 AA; 9393 MW; 63128984CBAF0531 CRC64;

Query Match 2.3%; Score 8; DB 1; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 0.86;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 305 LSIRKTLTP 312  
 |||||  
 Db 22 LSIRKTLTP 29

RESULT 6  
 V187\_BPT7 STANDARD; PRT; 83 AA.

AC P03788;  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 01-MAR-1989 (Rel. 10, Last annotation update)  
 DE Gene 18.7 protein.  
 GN 18.7.  
 OS Bacteriophage T7.  
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;  
 CC T7-like phages.  
 RX NCBI\_TaxID=10760;

CC T7-like phages.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=83241725; PubMed=6864790;  
 RA Dunn J.J., Studier F.W.;  
 RT "Complete nucleotide sequence of bacteriophage T7 DNA and the  
 RT locations of T7 genetic elements."  
 RT J. Mol. Biol. 166:477-535(1983).  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

CC EMBL: V01146; CAA24439.1; -  
 DR PIR: A04413; WBBP73.  
 DR PIR: S42337; S42337.  
 SQ SEQUENCE 83 AA; 9326 MW; B3DB055DE79A53C5 CRC64;

Query Match 2.3%; Score 8; DB 1; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 0.86;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 305 LSIRKTLTP 312  
 |||||  
 Db 22 LSIRKTLTP 29

RESULT 7  
 YD18\_YEAST STANDARD; PRT; 396 AA.

AC Q12185;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Hypothetical 45.9 kDa protein in KCS1-GCV1 intergenic region.  
 GN YD018C OR Y09335.04C OR PZF396.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 CC [1]  
 RN RP SEQUENCE FROM N.A.

RC STRAIN-S288C / AB972;  
 RA Dedman K., Brown D., Hamlyn N., Bowman S., Barrell B.G.,  
 RA Rajandream M.A.;  
 RL Submitted (MAY-1995) to the EMBL/Genbank/DBJ databases.

RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97051598; PubMed=8896275;  
 RA Elde L.G., Sander C., Prydz H.;  
 RT "Sequencing and analysis of a 35.4 kb region on the right arm of  
 RT Chromosome IV from Saccharomyces cerevisiae reveal 23 open reading  
 RT frames."  
 RT Yeast 12:1085-1090(1996).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).  
 CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
 CC ACYLTRANSFERASE FAMILY.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

CC EMBL: Z74314; CAA98838.1; -  
 DR EMBL: X95966; CAA65210.1; -  
 DR EMBL: Z49770; CAA69843.1; -  
 DR SGD: S0002425; YDR018C.  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KW Hypothetical protein, phospholipid biosynthesis; Transferase;  
 KW Acyltransferase; Transmembrane.  
 FT TRANSMEM 27 47  
 FT TRANSMEM 69 89  
 FT TRANSMEM 123 143  
 FT TRANSMEM 372 392  
 FT POTENTIAL.  
 FT POTENTIAL.  
 SQ SEQUENCE 396 AA; 45938 MW; 9F55AB56C8BD44DD CRC64;

Query Match 2.3%; Score 8; DB 1; Length 396;  
 Best Local Similarity 100.0%; Pred. No. 3.2;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 211 LDATYDVT 218  
 |||||  
 Db 265 LDATYDVT 272

RESULT 8  
 HSCA\_BUCAI STANDARD; PRT; 611 AA.

AC P57660;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Chapterone protein hscA homolog.  
 GN HSCA OR BU605.  
 OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum  
 OS symbiotic bacterium).  
 CC Bacteria; Proteobacteria; gamma subdivision; Buchnera.  
 OX NCBI\_TaxID=118099;  
 CC [1]

RN RP SEQUENCE FROM N.A.  
 RC STRAIN-TOKYO 1998;  
 RX MEDLINE=20445173; PubMed=10993077;  
 RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;  
 RT "Genome sequence of the endocellular bacterial symbiont of aphids  
 RT Buchnera sp. Aps."  
 RL Nature 407:81-86(2000).  
 CC -1- FUNCTION: PROBABLE CHAPERONE. HAS A LOW INTRINSIC ATPASE ACTIVITY  
 CC WHICH IS MARKEDLY STIMULATED BY HSCB (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.  
 CC -----



CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AP001119; BAB13289.1;  
 DR InterPro: IPR001023; HSP70.  
 DR Pfam: PF00012; HSP70.1.  
 DR PRINTS: PR00301; HEATSHOCK70.  
 DR PROSITE: PS00297; HSP70.1; 1.  
 DR PROSITE: PS00329; HSP70.2; 1.  
 DR PROSITE: PS01036; HSP70.3; 1.  
 KW Chaperone: ATP-binding; Complete proteome.  
 SQ SEQUENCE 611 AA; 68837 MW; 32B567C3073082A CRC64;

Query Match 2.3%; Score 8; DB 1; Length 611;  
 Best Local Similarity 100.0%; Pred. No. 4.6;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RYLPSV 9  
 DB 42 RYLPSV 49

RESULT 9  
 ID YJDO\_ECOLI STANDARD; PRT; 57 AA;  
 AC P58038;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Hypothetical protein yjdo.  
 GN YJDO OR B4128.1 OR Z5731 OR EC55110.1.  
 OS Escherichia coli, and  
 OS Escherichia coli O157:H7.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 NX NCBI\_TaxId=562, 83334;  
 RP [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12 / MG1655;  
 RC MEDLINE-9742617; PubMed-9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12.";  
 RL Science 277:1453-1474(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-O157:H7 / EDL933 / ATCC 700927;  
 RC MEDLINE-21074935; PubMed-11206551;  
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
 RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,  
 RA Apodaca J., Nantharajan T.S., Lin J., Yen G., Schwartz D.C.,  
 RA Welch R.A., Blattner F.R.;  
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";  
 RL Nature 409:529-533(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-O157:H7 / RIMD 0509952;  
 RC MEDLINE-21156231; PubMed-11258796;  
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
 RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,  
 RA Kihara S., Shiba T., Hattori M., Shinagawa H.;  
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli

RT O157:H7 and genomic comparison with a laboratory strain K-12.";  
 RL DNA Res. 8:11-22(2001).  
 CC -1 SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1 SIMILARITY: STRONG, TO E.COLI YDCX.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AE000485; -: NOT\_ANNOTATED\_CDS.  
 DR EMBL; AE005646; AAG59328.1; -;  
 DR EMBL; AP002568; -: NOT\_ANNOTATED\_CDS.  
 DR EcoGene: EG14342; yjdo.  
 DR Hypothetical protein; Transmembrane; Complete proteome.  
 FT TRANSMEM 7 27  
 FT TRANSMEM 37 57  
 FT POTENTIAL.  
 SQ SEQUENCE 57 AA; 6555 MW; A3670A19500F75D6 CRC64;

Query Match 2.0%; Score 7; DB 1; Length 57;  
 Best Local Similarity 100.0%; Pred. No. 6.5;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25 RLISAF 31  
 DB 33 RLISAF 39

RESULT 10  
 ID PURI\_LACCA STANDARD; PRT; 194 AA.  
 AC P35853;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Amidophosphoribosyltransferase precursor (EC 2.4.2.14) (Glutamine  
 DE phosphoribosylpyrophosphate amidotransferase) (ATase) (GPATase)  
 DE (Fragment).  
 GN PURF.  
 OS Lactobacillus casei.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;  
 OC Lactobacillus.  
 NX NCBI\_TaxId=1562;  
 RP [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE-93012962; PubMed-1398079;  
 RA Gu Z.-M., Martindale D.W., Lee B.H.;  
 RT "Isolation and complete sequence of the purL gene encoding PGAM  
 RT synthase II in Lactobacillus casei.";  
 RL Gene 119:123-126(1992).  
 RN [2]  
 RP ERRATUM.  
 RC MEDLINE-94040790; PubMed-8224889;  
 RC Gu Z.-M., Martindale D.W., Lee B.H.;  
 RL Gene 133:147-147(1993).  
 CC -1 CATALYTIC ACTIVITY: 5-phospho-beta-D-ribose 1-diphosphate +  
 CC L-glutamate = L-glutamine + 5-phospho-alpha-D-ribose 1-diphosphate  
 CC + H(2)O.  
 CC -1 PATHWAY: FIRST STEP IN DE NOVO PURINE BIOSYNTHESIS.  
 CC -1 SIMILARITY: THE GATASE DOMAIN BELONGS TO TYPE-2 GLUTAMINE  
 CC AMIDOTRANSFERASES.  
 CC -1 SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE  
 CC PURINE/PYRIMIDINE PHOSPHORIBOSYLTRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: M65265; AAC36948.1; -.

DR PIR: PC1136; PC1136.

DR HSSP: P00497; 1A00.

DR MEROPS: C44.001; -.

DR InterPro: IPR000583; GATase\_2.

DR InterPro: IPR023375; Pur\_Pyr\_pr\_transf.

DR Pfam: PF00310; GATase\_2; 1. TRANSFER, PARTIAL.

DR PROSITE: PS00103; PUR\_PYR\_PR\_TRANSFER, PARTIAL.

DR PROSITE: PS00443; GATase\_TYPE\_II; 1.

KM Purine biosynthesis; Transferase; Glycosyltransferase.

FT PROPEP 1 11 BY SIMILARITY.

FT CHAIN 12 >194 AMIDOPHOSPHORIBOSYLTRANSFERASE.

FT ACT\_SITE 12 12 GATASE (INVOLVED IN SUBSTRATE BINDING AND CATALYSIS) (BY SIMILARITY).

FT NON\_TER 194 194

SO SEQUENCE 194 AA; 21144 MW; 4A788C8C5365D5EC CRC64;

Query Match 2.08; Score 7; DB 1; Length 194; Best Local Similarity 100.0%; Pred. No. 18; Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 100 ALGHVRY 106  
|111111|

DB 79. ALGHVRY 85

RESULT 11  
HBGF\_CERAE  
ID HBGF\_CERAE STANDARD; PRT; 208 AA.  
AC 009118:  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Heparin-binding EGF-like growth factor precursor (HB-EGF) (HBEGF)  
DE (Diphtheria toxin receptor) (DT-R).  
GN DTR OR HEGFL.  
OS Cercopithecus aethiops (Green monkey) (Grivet).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecidae;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecidae;  
NCBI\_Taxid:9534;  
[1]  
RN SEQUENCE FROM N.A.  
RP MEDLINE-92298386; PubMed-1606612;  
RA Naglich J.G.; Metherall J.E.; Russel D.W.; Eldels L.;  
RT "Expression cloning of a diphtheria toxin receptor: identity with a heparin-binding EGF-like growth factor precursor.";  
RL Cell 69:1051-1061(1992).  
[2]  
RN TOXIN-BINDING DOMAIN.  
RP MEDLINE-95126975; PubMed-7826391;  
RA Hooper K.P.; Eldels L.;  
RT "Localization of a critical diphtheria toxin-binding domain to the C-terminus of the mature heparin-binding EGF-like growth factor region of the diphtheria toxin receptor.";  
RL Biochem. Biophys. Res. Commun. 206:710-717(1995).  
[3]  
RN Proliferation; May be involved in macrophage-mediated cellular proliferation. It is mitogenic for fibroblasts and smooth muscle but not endothelial cells. It is able to bind EGF receptors with higher affinity than EGF itself and is a far more potent mitogen for smooth muscle cells than EGF (by similarity).  
CC SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MATURE HB-EGF IS RELEASED INTO THE EXTRACELLULAR SPACE AND PROBABLY BINDS TO A RECEPTOR (BY SIMILARITY).  
CC -1- PTM: O-GLYCOSYLATED (BY SIMILARITY).  
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: M93012; -; NOT\_ANNOTATED\_CDS.

DR HSSP: Q99075; 1XDT.

DR InterPro: IPR000561; EGF-like.

DR Pfam: PF00008; EGF; 1.

DR SMART: SM00181; EGF; 1.

DR PROSITE: PS00022; EGF\_1; 1.

DR PROSITE: PS01186; EGF\_2; 1.

KM Growth factor; Heparin-binding; Signal; EGF-like domain; Transmembrane; Glycoprotein; Receptor.

FT SIGNAL 1 19 POTENTIAL.

FT PROPEP 20 62 BY SIMILARITY.

FT CHAIN 63 148 HEPARIN-BINDING EGF-LIKE GROWTH FACTOR.

FT PROPEP 149 208 C-TERMINAL (POTENTIAL).

FT DOMAIN 20 160 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 161 184 POTENTIAL.

FT DOMAIN 185 208 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 104 144 EGF-LIKE.

FT CARBOHYD 75 75 O-LINKED (GALNAC. . .) (BY SIMILARITY).

FT CARBOHYD 85 85 O-LINKED (GALNAC. . .) (BY SIMILARITY).

FT DISULFID 108 121 BY SIMILARITY.

FT DISULFID 116 132 BY SIMILARITY.

FT DISULFID 134 143 BY SIMILARITY.

SO SEQUENCE 208 AA; 22985 MW; 8D108289A0485AE9 CRC64;

Query Match 2.08; Score 7; DB 1; Length 208; Best Local Similarity 100.0%; Pred. No. 19; Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 LLPSVYL 10  
|111111|

DB 3 LLPSVYL 9

RESULT 12  
HBGF\_HUMAN  
ID HBGF\_HUMAN STANDARD; PRT; 208 AA.  
AC 099075;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Heparin-binding EGF-like growth factor precursor (HB-EGF) (HBEGF)  
DE (Diphtheria toxin receptor) (DT-R).  
GN DTR OR HEGFL.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiindae; Homo.  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiindae; Homo.  
NCBI\_Taxid:9606;  
[1]  
RN SEQUENCE FROM N.A. AND SEQUENCE OF 73-93.  
RP TISSUE-Macrophage;  
RC MEDLINE-91157008; PubMed-1840698;  
RA Higashiyama S.; Abraham J.A.; Miller J.; Fiddes J.C.; Klagsbrun M.;  
RT "A heparin-binding growth factor secreted by macrophage-like cells that is related to EGF.";  
RL Science 251:936-939(1991).  
[2]  
RN SEQUENCE FROM N.A.  
RP Klumetly W.; Bondoc M.; Cheng J.; Connolly K.S.; Gunning K.M.; Kadner K.; Miguel T.; Miller C.; Ptluck S.; Pollard M.; Rojeski H.; Subramanian S.; Martin C.H.;  
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
[3]  
RN SEQUENCE OF 63-141 AND 143-148, AND CARBOHYDRATE-LINKAGE STTES.  
RP TISSUE-Histiocytic lymphoma;  
RC MEDLINE-92210596; PubMed-1556128;  
RA Higashiyama S.; Lau K.; Besner G.E.; Abraham J.A.; Klagsbrun M.;

RT "Structure of heparin-binding EGF-like growth factor. Multiple forms,  
 RT primary structure, and glycosylation of the mature protein.";  
 RL J. Biol. Chem. 267:6205-6212(1992).  
 RN [4]  
 RX TOXIN-BINDING DOMAIN.  
 RX MEDLINE-95138082; PubMed-7836353;  
 RA Mitamura T., Higashiyama S., Taniguchi N., Klagsbrun M., Mekada E.;  
 RT "Diphtheria toxin binds to the epidermal growth factor (EGF)-like  
 RT domain of human heparin-binding EGF-like growth factor/diphtheria  
 RT toxin receptor and inhibits specifically its mitogenic activity.";  
 RL J. Biol. Chem. 270:1015-1019(1995).  
 RN [5]  
 RP X-RAY CRYSTALLOGRAPHY (2.65 ANGSTROMS) OF 73-147 IN COMPLEX WITH TOX.  
 RA MEDLINE-98324089; PubMed-9659904;  
 RA Louis G.V., Yang W., Bowman M.E., Choe S.;  
 RT "Crystal structure of the complex of diphtheria toxin with an  
 RT extracellular fragment of its receptor.";  
 RL Mol. Cell 1:67-78(1997).  
 CC -1- FUNCTION: MAY BE INVOLVED IN MACROPHAGE-MEDIATED CELLULAR  
 CC PROLIFERATION. IT IS MITOGENIC FOR FIBROBLASTS AND SMOOTH MUSCLE  
 CC BUT NOT ENDOTHELIAL CELLS. IT IS ABLE TO BIND EGF RECEPTORS WITH  
 CC HIGHER AFFINITY THAN EGF ITSELF AND IS A FAR MORE POTENT MITOGEN  
 CC FOR SMOOTH MUSCLE CELLS THAN EGF.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MATURE HB-EGF IS  
 CC RELEASED INTO THE EXTRACELLULAR SPACE AND PROBABLY BINDS TO A  
 CC RECEPTOR.  
 CC -1- PTM: SEVERAL N-TERMINI HAVE BEEN IDENTIFIED BY DIRECT SEQUENCING.  
 CC THE FORMS WITH N-TERMINI 63, 73 AND 74 HAVE BEEN TESTED AND FOUND  
 CC TO BE BIOLOGICALLY ACTIVE.  
 CC -1- PTM: O-GLYCAN ATTACHMENT SITES WERE DETERMINED BY EDMAN  
 CC DEGRADATION. O-GLYCANASE DIGEST SUGGESTS MUCIN-TYPE GLYCOSYLATION  
 CC (DONE IN HB-EGF PURIFIED FROM HISTIOCYTIC LYMPHOMA CELL LINE  
 CC U-937).  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: M60278; AAA35956.1; -  
 DR EMBL: AC004634; AAC15470.1; -  
 DR PIR: A38432; A38432.  
 DR PDB: 1XDT; 25-FEB-98.  
 DR MIM: 126150; -  
 DR InterPro: IPR000561; EGF-like.  
 DR Pfam: PF00008; EGF\_1.  
 DR SMART: SM00181; EGF\_1.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 KW Growth factor; Heparin-binding; Signal; EGF-like domain;  
 KW Transmembrane; Glycoprotein; Receptor; 3D-structure.  
 FT SIGNAL 1 19  
 FT PROPEP 20 62  
 FT CHAIN 63 148  
 FT PROPEP 149 208  
 FT DOMAIN 20 160  
 FT TRANSMEM 161 184  
 FT DOMAIN 185 208  
 FT DOMAIN 104 144  
 FT DISULFID 108 121  
 FT DISULFID 116 132  
 FT DISULFID 134 143  
 FT CARBOHYD 75 75  
 FT CARBOHYD 85 85  
 SQ SEQUENCE 208 AA; 23067 MW; 2CA3C9D1B8291B51 CR64;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 LIPSVYL 10  
 Db 3 LIPSVYL 9  
 RESULT 13  
 ID HBGF\_PIG STANDARD; PRT; 208 AA.  
 AC 001580;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DE 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Heparin-binding EGF-like growth factor precursor (HB-EGF) (HBEGF).  
 GN DTR OR HEGFL.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Pascall J.C.;  
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 99-182 FROM N.A.  
 RC STRAIN=LARGE WHITE; TISSUE=Heart;  
 RX MEDLINE-93075016; PubMed-1445231;  
 RA Vaughan T.J., Pascall J.C., Brown K.D.;  
 RT "Tissue distribution of mRNA for heparin-binding epidermal growth  
 RT factor";  
 RL Biochem. J. 267:681-684(1992).  
 CC -1- FUNCTION: MAY BE INVOLVED IN MACROPHAGE-MEDIATED CELLULAR  
 CC PROLIFERATION. IT IS MITOGENIC FOR FIBROBLASTS AND SMOOTH MUSCLE  
 CC BUT NOT ENDOTHELIAL CELLS. IT IS ABLE TO BIND EGF RECEPTORS WITH  
 CC HIGHER AFFINITY THAN EGF ITSELF AND IS A FAR MORE POTENT MITOGEN  
 CC FOR SMOOTH MUSCLE CELLS THAN EGF (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MATURE HB-EGF IS  
 CC RELEASED INTO THE EXTRACELLULAR SPACE AND PROBABLY BINDS TO A  
 CC RECEPTOR (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: MACROPHAGE, MIDBRAIN, CEREBELLUM,  
 CC HYPOTHALAMUS, CEREBRAL CORTEX, BUTYBORETHRAL GLAND, LUNG, HEART  
 CC VENTRICLE, KIDNEY, SKIN, PROSTATE, SEMINAL VESICLE, TESTIS; AT LOW  
 CC LEVELS IN LYMPH NODE, THYMUS, SPLEEN; NOT DETECTED IN PITUITARY,  
 CC OLFACTORY BULB, THYROID, DUODENUM, PANCREAS, LIVER, SUBMAXILLARY  
 CC GLAND.  
 CC -1- PTM: O-GLYCOSYLATED (PROBABLE).  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: Y15731; CAA75740.1; -  
 DR EMBL: X67295; CAA47709.1; -  
 DR HSSP: Q99075; 1XDT.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR001336; EGF\_1.  
 DR Pfam: PF00008; EGF\_1.  
 DR PRINTS: PR00009; EGFTEF.  
 DR SMART: SM00181; EGF\_1.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 KW Signal; Growth factor; Heparin-binding; EGF-like domain;  
 KW Transmembrane; Glycoprotein.  
 FT SIGNAL 1 23  
 FT PROPEP 24 62  
 FT CHAIN 63 148  
 FT PROPEP 149 208  
 C-TERMINAL (POTENTIAL).

Query Match 2.0%; Score 7; DB 1; Length 208;  
 Best Local Similarity 100.0%; Pred. No. 19;

```

FT DOMAIN 24 161 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 162 182 POTENTIAL.
FT DOMAIN 183 208 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 104 144 EGF-LIKE.
FT DISULFID 108 121 BY SIMILARITY.
FT DISULFID 116 132 BY SIMILARITY.
FT DISULFID 134 143 BY SIMILARITY.
FT CARBOHYD 85 85 O-LINKED (GALNAC...) (BY SIMILARITY).
SQ SEQUENCE 208 AA; 22986 MW; 0A7DA97AE30C8967 CRC64;

Query Match 2.0%; Score 7; DB 1; Length 208;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 4 LPSVVL 10
Db 3 LPSVVL 9

RESULT 14
HBGF_RAT
ID HBGF_RAT STANDARD; PRT; 208 AA.
AC 006175;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Heparin-binding EGF-like growth factor precursor (HB-EGF) (HBEGF).
GN DFR OR HEGFL.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Macrophage;
RX MEDLINE=93135756; Pubmed=7678488;
RA Abraham J.A., Damm D., Bajardi A., Miller J., Klagsbrun M.,
RA Ezekowitz R.A.B.;
RT "Heparin-binding EGF-like growth factor: characterization of rat and
RT mouse cDNA clones, protein domain conservation across species, and
RT transcript expression in tissues.";
RL Biochem. Biophys. Res. Commun. 190:125-133(1993)
CC -1- FUNCTION: MAY BE INVOLVED IN MACROPHAGE-MEDIATED CELLULAR
CC PROLIFERATION. IT IS MITOGENIC FOR FIBROBLASTS AND SMOOTH MUSCLE
CC BUT NOT ENDOTHELIAL CELLS. IT IS ABLE TO BIND EGF RECEPTORS WITH
CC HIGHER AFFINITY THAN EGF ITSELF AND IS A FAR MORE POTENT MITOGEN
CC FOR SMOOTH MUSCLE CELLS THAN EGF.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MATURE HB-EGF IS
CC RELEASED INTO THE EXTRACELLULAR SPACE AND PROBABLY BINDS TO A
CC RECEPTOR.
CC -1- TISSUE SPECIFICITY: MOST ABUNDANT IN SKELETAL MUSCLE, LUNG, SPLEEN
CC BRAIN AND HEART.
CC -1- PTM: O-GLYCOSYLATED (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L05489; AAA81780.1; -.
CC PIR: JCI409; JCI409.
CC HSSP: Q99075; IADT.
CC InterPro: IPR000561; EGF-like.
CC Pfam: PF00008; EGF; 1.
CC SMART: SM00181; EGF; 1.
CC PROSITE: PS00022; EGF_1; 1.
CC PROSITE: PS01186; EGF_2; 1.
CC Signal; Growth factor; Heparin-binding; EGF-like domain;

```

```

KW Transmembrane; Glycoprotein.
FT SIGNAL 1 23 POTENTIAL.
FT PROPEP 24 62 BY SIMILARITY.
FT CHAIN 63 148 HEPARIN-BINDING EGF-LIKE GROWTH FACTOR.
FT PROPEP 149 208 C-TERMINAL (POTENTIAL).
FT DOMAIN 24 160 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 161 184 POTENTIAL.
FT DOMAIN 185 208 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 104 144 EGF-LIKE.
FT DISULFID 108 121 BY SIMILARITY.
FT DISULFID 116 132 BY SIMILARITY.
FT DISULFID 134 143 BY SIMILARITY.
FT CARBOHYD 85 85 O-LINKED (GALNAC...) (BY SIMILARITY).
SQ SEQUENCE 208 AA; 22843 MW; DDB0045E116D064C CRC64;

Query Match 2.0%; Score 7; DB 1; Length 208;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 4 LPSVVL 10
Db 3 LPSVVL 9

RESULT 15
HBX4_MOUSE
ID HBX4_MOUSE STANDARD; PRT; 250 AA.
AC P10284;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Homeobox protein Hox-B4 (Hox-2.6).
GN HOXB4 OR HOXB-4 OR HOX-2.6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89091992; Pubmed=2463210;
RA Graham A., Papalopulu N., Lorimer J., McVey J.H., Tuddenham E.G.D.,
RA Krangel R.;
RT "Characterization of a murine homeo box gene, Hox-2.6, related to the
RT Drosophila Deformed gene.";
RL Genes Dev. 2:1424-1438(1988).
CC -1- FUNCTION: SEQUENCE-SPECIFIC TRANSCRIPTION FACTOR WHICH IS PART OF
CC A DEVELOPMENTAL REGULATORY SYSTEM THAT PROVIDES CELLS WITH
CC SPECIFIC POSITIONAL IDENTITIES ON THE ANTERIOR-POSTERIOR AXIS.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: BELONGS TO THE ANTP FAMILY OF HOMEBOX PROTEINS.
CC "DEFORMED" SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M36654; AAA37848.1; -.
CC PIR: A31757; A31757.
CC HSSP: P02833; 9ANT.
CC TRANSFAC: T01728; -.
CC MGD: MGI:96185; Hoxb4.
CC InterPro: IPR001827; Antennapedia.
CC InterPro: IPR001356; Homeobox.
CC Pfam: PF00046; homeobox; 1.
CC PRINTS: PR00025; ANTENNAPEDIA.
CC PRINTS: PR00024; HOMEBOX.
CC SMART: SM00389; HOX; 1.
CC PROSITE: PS00027; HOMEBOX_1; 1.

```

DR PROSITE; PS00032; ANTENNAPEDIA; 1.  
 DR PROSITE; PS50071; HOMEBOX\_2; 1.  
 KW Homebox; DNA-binding; Developmental protein; Nuclear protein;  
 KW Transcription regulation.  
 FT DOMAIN 15 138 PRO-RICH, PART OF THE TRANSCRIPTIONAL  
 FT DOMAIN 71 86 ACTIVATION DOMAIN.  
 FT DOMAIN 140 145 POLY-PRO.  
 FT DNA\_BIND 161 220 ANTP-TYPE HEXAPEPTIDE.  
 SQ SEQUENCE 250 AA; 27519 MW; D09D477A0E585BE6 CRC64;

Query Match 2.0%; Score 7; DB 1; Length 250;  
 Best local Similarity 100.0%; Pred. No. 22;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 227 GGORRES 233  
 Db 40 GGORRES 46

Search completed: August 28, 2002, 11:27:18  
 Job time: 511 sec

... PAGE BLANK (USPTO)

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:08:11 ; Search time 31 Seconds  
(without alignments)  
179.646 Million cell updates/sec

Title: US-09-853-526-70  
Perfect score: 1203  
Sequence: 1 MRLLPSVVLGTAPRYVLA.....NYLDATYDVTVYEGKDDGG 228

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/Dackfilest1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1203	100.0	228	4	US-09-338-907-70 Sequence 70, Appl
2	1203	100.0	228	4	US-09-218-207-70 Sequence 70, Appl
3	1203	100.0	353	2	US-08-996-306-4 Sequence 4, Appl
4	1203	100.0	353	4	US-09-338-907-4 Sequence 4, Appl
5	1203	100.0	353	4	US-09-218-207-4 Sequence 4, Appl
6	1203	100.0	364	2	US-08-996-306-5 Sequence 5, Appl
7	1203	100.0	364	4	US-09-338-907-5 Sequence 5, Appl
8	1203	100.0	364	4	US-09-218-207-5 Sequence 5, Appl
9	1040	86.5	354	4	US-09-338-907-74 Sequence 74, Appl
10	1040	86.5	354	4	US-09-218-207-74 Sequence 74, Appl
11	976	81.1	185	4	US-09-338-907-136 Sequence 136, App
12	976	81.1	185	4	US-09-218-207-136 Sequence 136, App
13	976	81.1	315	4	US-09-338-907-134 Sequence 134, App
14	976	81.1	315	4	US-09-218-207-134 Sequence 134, App
15	970	80.6	300	4	US-09-338-907-135 Sequence 135, App
16	970	80.6	300	4	US-09-218-207-135 Sequence 135, App
17	830	69.0	291	4	US-09-338-907-127 Sequence 127, App
18	830	69.0	291	4	US-09-218-207-127 Sequence 127, App
19	820	68.2	182	4	US-09-338-907-133 Sequence 133, App
20	820	68.2	182	4	US-09-218-207-133 Sequence 133, App
21	657	54.6	261	4	US-09-338-907-128 Sequence 128, App
22	657	54.6	261	4	US-09-218-207-128 Sequence 128, App
23	597	49.6	238	4	US-09-338-907-126 Sequence 126, App
24	597	49.6	238	4	US-09-218-207-126 Sequence 126, App
25	449	37.3	97	4	US-09-338-907-132 Sequence 132, App
26	449	37.3	97	4	US-09-218-207-132 Sequence 132, App
27	332	27.6	77	4	US-09-338-907-125 Sequence 125, App

28	332	27.6	77	4	US-09-218-207-125	Sequence 125, App
29	330	27.4	66	4	US-09-338-907-131	Sequence 131, App
30	330	27.4	66	4	US-09-218-207-131	Sequence 131, App
31	329	27.3	68	4	US-09-338-907-130	Sequence 130, App
32	329	27.3	68	4	US-09-218-207-130	Sequence 130, App
33	329	27.3	90	4	US-09-338-907-129	Sequence 129, App
34	329	27.3	90	4	US-09-218-207-129	Sequence 129, App
35	269	22.4	374	2	US-08-454-267-2	Sequence 2, Appl
36	269	22.4	374	2	US-08-454-267-6	Sequence 2, Appl
37	269	22.4	374	2	US-08-941-319-2	Sequence 2, Appl
38	269	22.4	374	2	US-08-941-319-6	Sequence 6, Appl
39	269	22.4	374	4	US-09-035-098-2	Sequence 2, Appl
40	269	22.4	374	4	US-09-035-098-6	Sequence 6, Appl
41	269	22.4	374	4	US-09-215-252-5	Sequence 5, Appl
42	196	16.3	378	4	US-09-215-252-17	Sequence 17, Appl
43	187.5	15.6	376	4	US-09-215-252-13	Sequence 13, Appl
44	175.5	14.6	314	4	US-09-215-252-15	Sequence 15, Appl
45	171	14.2	295	2	US-08-454-267-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1  
US-09-338-907-70  
; Sequence 70, Application US/09338907  
; Patent No. 6265546  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Ilyu, Chumakov  
; APPLICANT: Boujelert, Lydie  
; TITLE OF INVENTION: PROSTATE CANCER GENE  
; FILE REFERENCE: GENSET 18CPICP  
; CURRENT APPLICATION NUMBER: US/09/338, 907  
; CURRENT FILING DATE: 1999-06-23  
; EARLIER APPLICATION NUMBER: 08/996, 306  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: 60/099, 658  
; EARLIER FILING DATE: 1998-09-09  
; EARLIER APPLICATION NUMBER: 09/218, 207  
; EARLIER FILING DATE: 1998-12-22  
; NUMBER OF SEQ ID NOS: 578  
; SOFTWARE: Patent.pm  
; SEQ ID NO 70  
; LENGTH: 228  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-338-907-70

Query Match	100.0%	Score 1203;	DB 4;	Length 228;
Best Local Similarity	100.0%;	Pred. No. 2e-126;		
Matches 228;	Conservative	0;	Mismatches 0;	Indels 0;
			Gaps 0;	
OY	1	MRLLPSVVLGTAPRYVLA	NGVWRLSAFLPARFQALDDRLRYCVYQSMVLEFFENYTG	60
DB	1	MRLLPSVVLGTAPRYVLA	NGVWRLSAFLPARFQALDDRLRYCVYQSMVLEFFENYTG	60
OY	61	VOILLTGDPKKNENIYLANH	OSTVDWIVADILAIRONALGHVRYLKEGKWLPLYGC	120
DB	61	VOILLTGDPKKNENIYLANH	OSTVDWIVADILAIRONALGHVRYLKEGKWLPLYGC	120
OY	121	YFAOHGIVYKSAKNEKEMR	KLOSVDAGCPMVLVFFPESTRNPNQOTVLSASQAF	180
DB	121	YFAOHGIVYKSAKNEKEMR	KLOSVDAGCPMVLVFFPESTRNPNQOTVLSASQAF	180
OY	181	AAORGIAVLKHLVTPRIKAT	HAFCMKNVLAIVTVTVYVYEGKDDGG	228
DB	181	AAORGIAVLKHLVTPRIKAT	HAFCMKNVLAIVTVTVYVYEGKDDGG	228
RESULT	2			

US-09-218-207-70  
; Sequence 70, Application US/09218207  
; Patent No. 6346381  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Ilya, Chumakov  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: Prostate cancer gene  
; FILE REFERENCE: GENSET.018CPI  
; CURRENT APPLICATION NUMBER: US/09/218,207  
; EARLIER FILING DATE: 1998-12-22  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: 60/099,658  
; EARLIER FILING DATE: 1998-09-09  
; NUMBER OF SEQ ID NOS: 578  
; SOFTWARE: Patent.pm  
; SEQ ID NO 70  
; LENGTH: 228  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-218-207-70

Query Match 100.0%; Score 1203; DB 4; Length 228;  
Best Local Similarity 100.0%; Pred. No. 2e-126;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRYLPSVVLGTAPTVVLAMGVRLLSATLPAPFQALDDRLCYCVQSVLFPFEVYTG 60  
Db 1 MRYLPSVVLGTAPTVVLAMGVRLLSATLPAPFQALDDRLCYCVQSVLFPFEVYTG 60  
OY 61 VOILLYGDLPRKNNIYILANHOSTVDMIVADILAIQNALGHVRYLKGSLKMLPLYGC 120  
Db 61 VOILLYGDLPRKNNIYILANHOSTVDMIVADILAIQNALGHVRYLKGSLKMLPLYGC 120  
OY 121 YFAOHGCIYVKRSKAKFNKRNKLOSVDAGTPMYLVIFPEGTRVPEOTKVLASQAF 180  
Db 121 YFAOHGCIYVKRSKAKFNKRNKLOSVDAGTPMYLVIFPEGTRVPEOTKVLASQAF 180  
OY 181 AAOAGLAVLKHVLTPIKATHVAFDCKMKNYLDIYDTVYVEGKDDGG 228  
Db 181 AAOAGLAVLKHVLTPIKATHVAFDCKMKNYLDIYDTVYVEGKDDGG 228

RESULT 3  
US-08-996-306-4  
; Sequence 4, Application US/08996306  
; Patent No. 5945522  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Chumakov, Ilya  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: Prostate cancer gene  
; NUMBER OF SEQUENCES: 68  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Knobe, Martens, Olson & Bear  
; STREET: 501 West Broadway  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92101-3505  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: Win95  
; SOFTWARE: Word  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/996,306  
; FILING DATE:  
; CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:  
; NAME: Israelsen, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: GENSET.018A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 353 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: SINGLE  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: PROTEIN  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: potential Transmembrane helix  
; LOCATION: 1..33  
; IDENTIFICATION METHOD: Rao and Argos method  
; FEATURE:  
; NAME/KEY: potential Transmembrane helix  
; LOCATION: 4..20  
; IDENTIFICATION METHOD: Klein, Kanehisa and Delisi method  
; FEATURE:  
; NAME/KEY: potential Transmembrane helix  
; LOCATION: 4..24  
; IDENTIFICATION METHOD: Eisenberg,Schwarz,Komarovy and Wall method  
; FEATURE:  
; NAME/KEY: potential N-myristoylation site  
; LOCATION: 12  
; IDENTIFICATION METHOD: prosite match  
; FEATURE:  
; NAME/KEY: potential Transmembrane helix  
; LOCATION: 50..70  
; IDENTIFICATION METHOD: Eisenberg,Schwarz,Komarovy and Wall method  
; FEATURE:  
; NAME/KEY: potential N-glycosylation site  
; LOCATION: 57  
; IDENTIFICATION METHOD: prosite match  
; FEATURE:  
; NAME/KEY: potential Transmembrane helix  
; LOCATION: 76..96  
; IDENTIFICATION METHOD: Eisenberg,Schwarz,Komarovy and Wall method  
; FEATURE:  
; NAME/KEY: potential Tyrosine kinase phosphorylation site  
; LOCATION: 78  
; IDENTIFICATION METHOD: prosite match  
; FEATURE:  
; NAME/KEY: potential Casein kinase II phosphorylation site  
; LOCATION: 84  
; IDENTIFICATION METHOD: prosite match  
; FEATURE:  
; NAME/KEY: potential Leucine zipper pattern  
; LOCATION: 94..115  
; IDENTIFICATION METHOD: prosite match  
; FEATURE:  
; NAME/KEY: potential N-myristoylation site  
; LOCATION: 119  
; IDENTIFICATION METHOD: prosite match  
; FEATURE:  
; NAME/KEY: potential Protein kinase C phosphorylation site  
; LOCATION: 133  
; IDENTIFICATION METHOD: prosite match  
; FEATURE:  
; NAME/KEY: potential Casein kinase II phosphorylation site  
; LOCATION: 147  
; IDENTIFICATION METHOD: prosite match  
; FEATURE:  
; NAME/KEY: potential Protein kinase C phosphorylation site  
; LOCATION: 194  
; IDENTIFICATION METHOD: prosite match



```
FEATURE:
NAME/KEY: potential Tyrosine kinase phosphorylation site
LOCATION: 215
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Tyrosine sulfatation site
LOCATION: 221
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential cAMP- and cGMP-dependent protein kinase phosphorylation site
LOCATION: 233
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Casein kinase II phosphorylation site
LOCATION: 235
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Protein kinase C phosphorylation site
LOCATION: 306
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Transmembrane helix
LOCATION: 310..330
IDENTIFICATION METHOD: Eisenberg,Schwarz,Komarony and Wall method
FEATURE:
NAME/KEY: potential N-myristoylation site
LOCATION: 319
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential N-myristoylation site
LOCATION: 323
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Amidation site
LOCATION: 329
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Transmembrane helix
LOCATION: 333..353
IDENTIFICATION METHOD: Eisenberg,Schwarz,Komarony and Wall method
FEATURE:
NAME/KEY: potential N-myristoylation site
LOCATION: 341
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Protein kinase C phosphorylation site
LOCATION: 350
IDENTIFICATION METHOD: prosite match
US-08-996-306-4
```

```
Query Match          100.0%; Score 1203; DB 2; Length 353;
Best Local Similarity 100.0%; Pred. No. 3,7e-126;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 MRYLLPSVVLCTAPTYVLAAGVWRLLSAFLLPARFYQALDDRLCYCYOSMVLFFPENYTG 60
DB 1 MRYLLPSVVLCTAPTYVLAAGVWRLLSAFLLPARFYQALDDRLCYCYOSMVLFFPENYTG 60
OY 61 VOILLYGDLPKKNENIYLANHQSIVDWIADILAIROMALGHVRYVLEKGLKWLPLYGC 120
DB 61 VOILLYGDLPKKNENIYLANHQSIVDWIADILAIROMALGHVRYVLEKGLKWLPLYGC 120
OY 121 YFAHGGTYVRSKSAFNEKEMRNKLOSYYVDAGTAPMVLVFFPGSTRNPEQTYVLSASQAF 180
DB 121 YFAHGGTYVRSKSAFNEKEMRNKLOSYYVDAGTAPMVLVFFPGSTRNPEQTYVLSASQAF 180
OY 181 AAGRLAVLKHVLPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDGG 228
DB 181 AAGRLAVLKHVLPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDGG 228
```

RESULT 4

```
US-09-338-907-4
Sequence 4, Application US/09338907
Patent No. 6265546
GENERAL INFORMATION:
APPLICANT: Cohen, Daniel
APPLICANT: Blumenfeld, Marta
APPLICANT: Ilya, Chumakov
APPLICANT: Bougueleret, Lydie
TITLE OF INVENTION: PROSTATE CANCER GENE
FILE REFERENCE: GENSET 18CPIGP
CURRENT APPLICATION NUMBER: US/09/338,907
CURRENT FILING DATE: 1999-06-23
EARLIER APPLICATION NUMBER: 08/996,306
EARLIER FILING DATE: 1997-12-22
EARLIER APPLICATION NUMBER: 60/099,658
EARLIER FILING DATE: 1998-09-09
EARLIER APPLICATION NUMBER: 09/218,207
EARLIER FILING DATE: 1998-12-22
NUMBER OF SEQ ID NOS: 578
SOFTWARE: Patent.pm
SEQ ID NO 4
LENGTH: 353
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: HELIX
LOCATION: 1..33
OTHER INFORMATION: Rao and Argos identification method, potential helix
FEATURE:
NAME/KEY: HELIX
LOCATION: 4..20
OTHER INFORMATION: Klein, Kanehisa and Delisl identification method, potent
FEATURE:
NAME/KEY: HELIX
LOCATION: 4..24
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
FEATURE:
NAME/KEY: MYRISTATE
LOCATION: 12..16
OTHER INFORMATION: Prosite match
FEATURE:
NAME/KEY: HELIX
LOCATION: 50..70
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
FEATURE:
NAME/KEY: CARBOHYD
LOCATION: 57..59
OTHER INFORMATION: Prosite match
FEATURE:
NAME/KEY: HELIX
LOCATION: 76..96
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
FEATURE:
NAME/KEY: PHOSPHORYLATION
LOCATION: 78
OTHER INFORMATION: potential Tyrosine kinase site, Prosite match
FEATURE:
NAME/KEY: PHOSPHORYLATION
LOCATION: 84
OTHER INFORMATION: potential Tyrosine kinase site, Prosite match
FEATURE:
NAME/KEY: SITE
LOCATION: 94..115
OTHER INFORMATION: potential leucine zipper site, Prosite match
FEATURE:
NAME/KEY: MYRISTATE
LOCATION: 119..123
OTHER INFORMATION: potential site, Prosite match
FEATURE:
NAME/KEY: PHOSPHORYLATION
LOCATION: 133
OTHER INFORMATION: potential protein kinase C, Prosite match
FEATURE:
```

```

? NAME/KEY: PHOSPHORYLATION
? LOCATION: 147
? OTHER INFORMATION: potential caseine kinase II site, Prosite match
? FEATURE:
? NAME/KEY: PHOSPHORYLATION
? LOCATION: 194
? OTHER INFORMATION: potential protein kinase C, Prosite match
? FEATURE:
? NAME/KEY: PHOSPHORYLATION
? LOCATION: 215
? OTHER INFORMATION: potential Tyrosine kinase site, Prosite match
? FEATURE:
? NAME/KEY: SULFATATION
? LOCATION: 221
? OTHER INFORMATION: Prosite match
? FEATURE:
? NAME/KEY: PHOSPHORYLATION
? LOCATION: 233
? OTHER INFORMATION: potential cAMP and cGMP dependant protein kinase site, Pr
? FEATURE:
? NAME/KEY: PHOSPHORYLATION
? LOCATION: 235
? OTHER INFORMATION: potential caseine kinase II site, Prosite match
? FEATURE:
? NAME/KEY: PHOSPHORYLATION
? LOCATION: 306
? OTHER INFORMATION: potential protein kinase C, Prosite match
? FEATURE:
? NAME/KEY: HELIX
? LOCATION: 310..330
? OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall Identification method, F
? FEATURE:
? NAME/KEY: MYRISTATE
? LOCATION: 319..323
? OTHER INFORMATION: Prosite match
? FEATURE:
? NAME/KEY: MYRISTATE
? LOCATION: 323..327
? OTHER INFORMATION: Prosite match
? FEATURE:
? NAME/KEY: AMIDATION
? LOCATION: 329
? OTHER INFORMATION: Prosite match
? FEATURE:
? NAME/KEY: HELIX
? LOCATION: 333..353
? OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall Identification method, F
? FEATURE:
? NAME/KEY: MYRISTATE
? LOCATION: 341..345
? OTHER INFORMATION: Prosite match
? FEATURE:
? NAME/KEY: PHOSPHORYLATION
? LOCATION: 350
? OTHER INFORMATION: potential protein kinase C, Prosite match
US-09-338-907-4
```

Query Match 100.0%; Score 1203; DB 4; Length 353;  
Best Local Similarity 100.0%; Pred. No. 3.7e-126;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 MRYLLPSVVLGTAPTYVLAWGVRLSAFLPARFYQALDDRLCYVQSMVLPFFENVYTG 60
    |||||
Db 1 MRYLLPSVVLGTAPTYVLAWGVRLSAFLPARFYQALDDRLCYVQSMVLPFFENVYTG 60
Qy 61 VQILLYDGLPKNKENITYLANHSTVDWIVADILAIROMALGHVRYVLKESLWLPPLYGC 120
    |||||
Db 61 VQILLYDGLPKNKENITYLANHSTVDWIVADILAIROMALGHVRYVLKESLWLPPLYGC 120
Qy 121 YFAOHGSIYVRSKAKFEKEMKRNKLOSVDAGTPMYLVIPPEGRYRNPEDQTKVLSASQAF 180
    |||||
Db 121 YFAOHGSIYVRSKAKFEKEMKRNKLOSVDAGTPMYLVIPPEGRYRNPEDQTKVLSASQAF 180
```

```

Qy 181 AAGRLAVLKHVLPRIKATHVAFDCMKNYLDATYDVTVYVEGKDDG 228
    |||||
Db 181 AAGRLAVLKHVLPRIKATHVAFDCMKNYLDATYDVTVYVEGKDDG 228
```

```

RESULT 5
US-09-218-207-4
? Sequence 4, Application US/09218207
? Patent No. 6346381
? GENERAL INFORMATION:
? APPLICANT: Cohen, Daniel
? APPLICANT: Blumenfeld, Marta
? APPLICANT: Ilya, Chumakov
? APPLICANT: Bougueleret, Lydie
? TITLE OF INVENTION: Prostate cancer gene
? FILE REFERENCE: GENSET 018CPI
? CURRENT APPLICATION NUMBER: US/09/218, 207
? EARLIER FILING DATE: 1998-12-22
? EARLIER APPLICATION NUMBER: 08/996, 306
? EARLIER FILING DATE: 1997-12-22
? EARLIER APPLICATION NUMBER: 60/099, 658
? NUMBER OF SEQ ID NOS: 578
? SOFTWARE: PatentL.pm
? SEQ ID NO 4
? LENGTH: 353
? TYPE: PRT
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: HELIX
? LOCATION: 1..33
? OTHER INFORMATION: Rao and Argos Identification method, potential helix
? FEATURE:
? NAME/KEY: HELIX
? LOCATION: 4..20
? OTHER INFORMATION: Klein, Kanehisa and Delist Identification method, potent
? FEATURE:
? NAME/KEY: HELIX
? LOCATION: 4..24
? OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall Identification method,
? FEATURE:
? NAME/KEY: MYRISTATE
? LOCATION: 12..16
? OTHER INFORMATION: Prosite match
? FEATURE:
? NAME/KEY: HELIX
? LOCATION: 50..70
? OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall Identification method,
? FEATURE:
? NAME/KEY: CARBOHYD
? LOCATION: 57..59
? OTHER INFORMATION: Prosite match
? FEATURE:
? NAME/KEY: HELIX
? LOCATION: 76..96
? OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall Identification method,
? FEATURE:
? NAME/KEY: PHOSPHORYLATION
? LOCATION: 78
? OTHER INFORMATION: potential Tyrosine kinase site, Prosite match
? FEATURE:
? NAME/KEY: PHOSPHORYLATION
? LOCATION: 84
? OTHER INFORMATION: potential caseine kinase II site, Prosite match
? FEATURE:
? NAME/KEY: SITE
? LOCATION: 94..115
? OTHER INFORMATION: potential leucine zipper site, Prosite match
? FEATURE:
? NAME/KEY: MYRISTATE
? LOCATION: 119..123
? OTHER INFORMATION: potential site, Prosite match
```

```

FEATURE:
NAME/KEY: PHOSPHORYLATION
LOCATION: 133
OTHER INFORMATION: potential protein kinase C, Prosite match
FEATURE:
NAME/KEY: PHOSPHORYLATION
LOCATION: 147
OTHER INFORMATION: potential caseine kinase II site, Prosite match
FEATURE:
NAME/KEY: PHOSPHORYLATION
LOCATION: 194
OTHER INFORMATION: potential protein kinase C, Prosite match
FEATURE:
NAME/KEY: PHOSPHORYLATION
LOCATION: 215
OTHER INFORMATION: potential tyrosine kinase site, Prosite match
FEATURE:
NAME/KEY: SULFATATION
LOCATION: 221
OTHER INFORMATION: Prosite match
FEATURE:
NAME/KEY: PHOSPHORYLATION
LOCATION: 233
OTHER INFORMATION: potential CAMP and cGMP dependant protein kinase site, Pr
FEATURE:
NAME/KEY: PHOSPHORYLATION
LOCATION: 235
OTHER INFORMATION: potential caseine kinase II site, Prosite match
FEATURE:
NAME/KEY: PHOSPHORYLATION
LOCATION: 306
OTHER INFORMATION: potential protein kinase C, Prosite match
FEATURE:
NAME/KEY: HELIX
LOCATION: 310..330
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F
FEATURE:
NAME/KEY: MYRISTATE
LOCATION: 319..323
OTHER INFORMATION: Prosite match
FEATURE:
NAME/KEY: MYRISTATE
LOCATION: 323..327
OTHER INFORMATION: Prosite match
FEATURE:
NAME/KEY: AMIDATION
LOCATION: 329
OTHER INFORMATION: Prosite match
FEATURE:
NAME/KEY: HELIX
LOCATION: 333..353
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F
FEATURE:
NAME/KEY: MYRISTATE
LOCATION: 341..345
OTHER INFORMATION: Prosite match
FEATURE:
NAME/KEY: PHOSPHORYLATION
LOCATION: 350
OTHER INFORMATION: potential protein kinase C, Prosite match
US-09-218-207-4
```

```

Query Match          100.0%; Score 1203; DB 4; Length 353;
Best Local Similarity 100.0%; Pred. No. 3.7e-126;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRLPLSVVLGTAPRYVLAAGVWRLLSAFLPARFYQALDDRLCYVQSMVLEFFENYTG 60
    |||
DB 1 MRLPLSVVLGTAPRYVLAAGVWRLLSAFLPARFYQALDDRLCYVQSMVLEFFENYTG 60
    |||
QY 61 VOILLGDLPRKNENIYLANHOSVDWIVADILAIRONALGHVRYVLEGLKWLPLYGC 120
    |||
```

```

DB 61 VOILLGDLPRKNENIYLANHOSVDWIVADILAIRONALGHVRYVLEGLKWLPLYGC 120
QY 121 YFAOHGIGYVRRSAKFNEMRNKLOSVDAGTPTLVTFPESTRNPEOTKVLASQA 180
    |||
DB 121 YFAOHGIGYVRRSAKFNEMRNKLOSVDAGTPTLVTFPESTRNPEOTKVLASQA 180
QY 181 AAGRGVLAHLVLTPTRIKATHVAFDCMKNYLDAIVDTVYVEGKDDGG 228
    |||
DB 181 AAGRGVLAHLVLTPTRIKATHVAFDCMKNYLDAIVDTVYVEGKDDGG 228
```

```

RESULT 6
US-08-996-306-5
; Sequence 5, Application US/08996306
; Patent No. 5945522
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Chumakov, Ilya
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; NUMBER OF SEQUENCES: 68
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbey, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/996,306
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelien, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: GENSET.018A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 364 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; ORGANISM: Homo sapiens
; ORIGINAL SOURCE:
; FEATURE:
; NAME/KEY: potential protein
; LOCATION: 1..364
US-08-996-306-5
```

```

Query Match          100.0%; Score 1203; DB 2; Length 364;
Best Local Similarity 100.0%; Pred. No. 3.9e-126;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRLPLSVVLGTAPRYVLAAGVWRLLSAFLPARFYQALDDRLCYVQSMVLEFFENYTG 60
    |||
DB 12 MRLPLSVVLGTAPRYVLAAGVWRLLSAFLPARFYQALDDRLCYVQSMVLEFFENYTG 71
    |||
QY 61 VOILLGDLPRKNENIYLANHOSVDWIVADILAIRONALGHVRYVLEGLKWLPLYGC 120
    |||
DB 72 VOILLGDLPRKNENIYLANHOSVDWIVADILAIRONALGHVRYVLEGLKWLPLYGC 131
    |||
QY 121 YFAOHGIGYVRRSAKFNEMRNKLOSVDAGTPTLVTFPESTRNPEOTKVLASQA 180
    |||
```

```
Db 132 YFAOHGIIYKRSKAKFNEKRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASQAF 191
Qy 181 AAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDG 228
Db 192 AAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDG 239
```

```
RESULT 7
US-09-338-907-5
; Sequence 5, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CP1CP
; CURRENT APPLICATION NUMBER: US/09/338,907
; EARLIER FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218,207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 5
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-338-907-5
```

```
Query Match 100.0%; Score 1203; DB 4; Length 364;
Best Local Similarity 100.0%; Pred. No. 3.9e-126;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRYLLPSVVLGTAPTYVLAAGVWRLSAPLPARFYQALDDRLCYCYQSWLFFPENYTG 60
Db 12 MRYLLPSVVLGTAPTYVLAAGVWRLSAPLPARFYQALDDRLCYCYQSWLFFPENYTG 71
Qy 61 VOILLYGDLPKRNKENITYLANHOSVDWIVADILAIROMALGHVRYVLEKGLKWLPLYGC 120
Db 72 VOILLYGDLPKRNKENITYLANHOSVDWIVADILAIROMALGHVRYVLEKGLKWLPLYGC 131
Qy 121 YFAOHGIIYKRSKAKFNEKRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASQAF 180
Db 132 YFAOHGIIYKRSKAKFNEKRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASQAF 191
Qy 181 AAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDG 228
Db 192 AAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDG 239
```

```
RESULT 8
US-09-218-207-5
; Sequence 5, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; FILE REFERENCE: GENSET.01BCP1
; CURRENT APPLICATION NUMBER: US/09/218,207
; EARLIER FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
```

```
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 5
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-218-207-5
```

```
Query Match 100.0%; Score 1203; DB 4; Length 364;
Best Local Similarity 100.0%; Pred. No. 3.9e-126;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRYLLPSVVLGTAPTYVLAAGVWRLSAPLPARFYQALDDRLCYCYQSWLFFPENYTG 60
Db 12 MRYLLPSVVLGTAPTYVLAAGVWRLSAPLPARFYQALDDRLCYCYQSWLFFPENYTG 71
Qy 61 VOILLYGDLPKRNKENITYLANHOSVDWIVADILAIROMALGHVRYVLEKGLKWLPLYGC 120
Db 72 VOILLYGDLPKRNKENITYLANHOSVDWIVADILAIROMALGHVRYVLEKGLKWLPLYGC 131
Qy 121 YFAOHGIIYKRSKAKFNEKRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASQAF 180
Db 132 YFAOHGIIYKRSKAKFNEKRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASQAF 191
Qy 181 AAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDG 228
Db 192 AAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDG 239
```

```
RESULT 9
US-09-338-907-74
; Sequence 74, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CP1CP
; CURRENT APPLICATION NUMBER: US/09/338,907
; EARLIER FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218,207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 74
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-338-907-74
```

```
Query Match 86.5%; Score 1040; DB 4; Length 354;
Best Local Similarity 85.9%; Pred. No. 6e-108;
Matches 195; Conservative 19; Mismatches 13; Indels 0; Gaps 0;

Qy 1 MRYLLPSVVLGTAPTYVLAAGVWRLSAPLPARFYQALDDRLCYCYQSWLFFPENYTG 60
Db 1 MRYLLPSVVLGTAPTYVLAAGVWRLSAPLPARFYQALDDRLCYCYQSWLFFPENYTG 60
Qy 61 VOILLYGDLPKRNKENITYLANHOSVDWIVADILAIROMALGHVRYVLEKGLKWLPLYGC 120
Db 61 VOILLYGDLPKRNKENITYLANHOSVDWIVADILAIROMALGHVRYVLEKGLKWLPLYGC 120
Qy 121 YFAOHGIIYKRSKAKFNEKRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASQAF 180
Db 121 YFAOHGIIYKRSKAKFNEKRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASQAF 180
```

Db 121 YFAQHGIIYVKRSKAFNDKEMSKLOSIVNAGTPTMYLTFPESTRYNATYTKLLSASQAF 180  
QY 181 AAORGLAVLKHVLTPTRIKATHVAFDCMKMYLDAIVDTVYVEGKDDG 227  
Db 181 AAORGLAVLKHVLTPTRIKATHVAFDSMKSHLDIYDTVYVEGKENG 227

RESULT 10  
US-09-218-207-74  
; Sequence 74, Application US/09218207  
; Patent No. 6346381  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Ilya, Chumakov  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: Prostate cancer gene  
; FILE REFERENCE: GENSET.018CPI  
; CURRENT APPLICATION NUMBER: US/09/218,207  
; CURRENT FILING DATE: 1998-12-22  
; EARLIER APPLICATION NUMBER: 08/996,306  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: 60/099,658  
; EARLIER FILING DATE: 1998-09-09  
; NUMBER OF SEQ ID NOS: 578  
; SOFTWARE: Patent.pm  
; SEQ ID NO 74  
; LENGTH: 354  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-218-207-74

Query Match 86.5%; Score 1040; DB 4; Length 354;  
Best Local Similarity 85.9%; Pred. No. 6e-108;  
Matches 195; Conservative 19; Mismatches 13; Indels 0; Gaps 0;

QY 1 MRYLPSVVLGTAPTYVLAAGVWRLSAPLPARFYQALDDRLCYQOSVLEFFENYTG 60  
Db 1 MRYLPSVVLGTAPTYVLAAGVWRLSAPLPARFYQALDDRLCYQOSVLEFFENYTG 60  
QY 61 VOILLYGDLPRKNENIITYLANHOSYVDWIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120  
Db 61 VOILLYGDLPRKNENIITYLANHOSYVDWIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120  
QY 121 YFAQHGIIYVKRSKAFNDKEMSKLOSIVNAGTPTMYLTFPESTRYNATYTKLLSASQAF 180  
Db 121 YFAQHGIIYVKRSKAFNDKEMSKLOSIVNAGTPTMYLTFPESTRYNATYTKLLSASQAF 180  
QY 181 AAORGLAVLKHVLTPTRIKATHVAFDCMKMYLDAIVDTVYVEGKDDG 227  
Db 181 AAORGLAVLKHVLTPTRIKATHVAFDSMKSHLDIYDTVYVEGKENG 227

RESULT 11  
US-09-338-907-136  
; Sequence 136, Application US/09338907  
; Patent No. 6265546  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Ilya, Chumakov  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: PROSTATE CANCER GENE  
; FILE REFERENCE: GENSET.18CPICP  
; CURRENT APPLICATION NUMBER: US/09/338,907  
; CURRENT FILING DATE: 1999-06-23  
; EARLIER APPLICATION NUMBER: 08/996,306  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: 60/099,658  
; EARLIER FILING DATE: 1998-09-09  
; EARLIER APPLICATION NUMBER: 09/218,207  
; EARLIER FILING DATE: 1998-12-22

; NUMBER OF SEQ ID NOS: 578  
; SOFTWARE: Patent.pm  
; SEQ ID NO 136  
; LENGTH: 185  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: 81..83  
; OTHER INFORMATION: Box I  
; NAME/KEY: SITE  
; LOCATION: 160..165  
; OTHER INFORMATION: Box II  
US-09-338-907-136

Query Match 81.1%; Score 976; DB 4; Length 185;  
Best Local Similarity 100.0%; Pred. No. 3.3e-101;  
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLPSVVLGTAPTYVLAAGVWRLSAPLPARFYQALDDRLCYQOSVLEFFENYTG 60  
Db 1 MRYLPSVVLGTAPTYVLAAGVWRLSAPLPARFYQALDDRLCYQOSVLEFFENYTG 60  
QY 61 VOILLYGDLPRKNENIITYLANHOSYVDWIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120  
Db 61 VOILLYGDLPRKNENIITYLANHOSYVDWIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120  
QY 121 YFAQHGIIYVKRSKAFNDKEMSKLOSIVNAGTPTMYLTFPESTRYNATYTKLLSASQAF 180  
Db 121 YFAQHGIIYVKRSKAFNDKEMSKLOSIVNAGTPTMYLTFPESTRYNATYTKLLSASQAF 180  
QY 181 AAORG 185  
Db 181 AAORG 185

RESULT 12  
US-09-218-207-136  
; Sequence 136, Application US/09218207  
; Patent No. 6346381  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Ilya, Chumakov  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: Prostate cancer gene  
; FILE REFERENCE: GENSET.018CPI  
; CURRENT APPLICATION NUMBER: US/09/218,207  
; CURRENT FILING DATE: 1998-12-22  
; EARLIER APPLICATION NUMBER: 08/996,306  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: 60/099,658  
; EARLIER FILING DATE: 1998-09-09  
; NUMBER OF SEQ ID NOS: 578  
; SOFTWARE: Patent.pm  
; SEQ ID NO 136  
; LENGTH: 185  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: 81..83  
; OTHER INFORMATION: Box I  
; NAME/KEY: SITE  
; LOCATION: 160..165  
; OTHER INFORMATION: Box II  
US-09-218-207-136

Query Match 81.1%; Score 976; DB 4; Length 185;

Best Local Similarity 100.0%; Pred. No. 3.3e-101;  
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQVQSWLFFFEYNTG 60  
|  
DB 1 MRYLLPSVLLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQVQSWLFFFEYNTG 60  
QY 61 VOILLYGDLPKKNENIITYLANHOSTVDMIVADILAIROMALGHVRYVLKEGLKWLPLYGC 120  
|  
DB 61 VOILLYGDLPKKNENIITYLANHOSTVDMIVADILAIROMALGHVRYVLKEGLKWLPLYGC 120  
QY 121 YFAOHGIIYVRSKAKFNEKEKRNKLOSVDAGTPMYLVIPEEGTRYNPEQTKVLSASQAF 180  
|  
DB 121 YFAOHGIIYVRSKAKFNEKEKRNKLOSVDAGTPMYLVIPEEGTRYNPEQTKVLSASQAF 180  
QY 181 AAORG 185  
|  
DB 181 AAORG 185

## RESULT 13

US-09-338-907-134  
; Sequence 134, Application US/09338907  
; Patent No. 6265546  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Ilya, Chumakov  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: PROSTATE CANCER GENE  
; FILE REFERENCE: GENSET.18CPLCP  
; CURRENT APPLICATION NUMBER: US/09/338.907  
; EARLIER FILING DATE: 1999-06-23  
; EARLIER APPLICATION NUMBER: 08/996,306  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: 60/099,658  
; EARLIER FILING DATE: 1998-09-09  
; EARLIER APPLICATION NUMBER: 09/218,207  
; EARLIER FILING DATE: 1998-12-22  
; NUMBER OF SEQ ID NOS: 578  
; SOFTWARE: Patent.pm  
; SEQ ID NO 134  
; LENGTH: 315  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; NAME/KEY: SITE  
; LOCATION: 81..83  
; OTHER INFORMATION: Box I  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: 160..165  
; OTHER INFORMATION: Box II  
US-09-338-907-134

Query Match 81.1%; Score 976; DB 4; Length 315;  
Best Local Similarity 100.0%; Pred. No. 7.1e-101;  
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQVQSWLFFFEYNTG 60  
|  
DB 1 MRYLLPSVLLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQVQSWLFFFEYNTG 60  
QY 61 VOILLYGDLPKKNENIITYLANHOSTVDMIVADILAIROMALGHVRYVLKEGLKWLPLYGC 120  
|  
DB 61 VOILLYGDLPKKNENIITYLANHOSTVDMIVADILAIROMALGHVRYVLKEGLKWLPLYGC 120  
QY 121 YFAOHGIIYVRSKAKFNEKEKRNKLOSVDAGTPMYLVIPEEGTRYNPEQTKVLSASQAF 180  
|  
DB 121 YFAOHGIIYVRSKAKFNEKEKRNKLOSVDAGTPMYLVIPEEGTRYNPEQTKVLSASQAF 180  
QY 181 AAORG 185  
|  
DB 181 AAORG 185

DB 181 AAORG 185  
|  
DB 181 AAORG 185

## RESULT 14

US-09-218-207-134  
; Sequence 134, Application US/09218207  
; Patent No. 6346381  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Ilya, Chumakov  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: Prostate cancer gene  
; FILE REFERENCE: GENSET.018CPL  
; CURRENT APPLICATION NUMBER: US/09/218,207  
; EARLIER FILING DATE: 1998-12-22  
; EARLIER APPLICATION NUMBER: 08/996,306  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: 60/099,658  
; EARLIER FILING DATE: 1998-09-09  
; NUMBER OF SEQ ID NOS: 578  
; SOFTWARE: Patent.pm  
; SEQ ID NO 134  
; LENGTH: 315  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; NAME/KEY: SITE  
; LOCATION: 81..83  
; OTHER INFORMATION: Box I  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: 160..165  
; OTHER INFORMATION: Box II  
US-09-218-207-134

Query Match 81.1%; Score 976; DB 4; Length 315;  
Best Local Similarity 100.0%; Pred. No. 7.1e-101;  
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQVQSWLFFFEYNTG 60  
|  
DB 1 MRYLLPSVLLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQVQSWLFFFEYNTG 60  
QY 61 VOILLYGDLPKKNENIITYLANHOSTVDMIVADILAIROMALGHVRYVLKEGLKWLPLYGC 120  
|  
DB 61 VOILLYGDLPKKNENIITYLANHOSTVDMIVADILAIROMALGHVRYVLKEGLKWLPLYGC 120  
QY 121 YFAOHGIIYVRSKAKFNEKEKRNKLOSVDAGTPMYLVIPEEGTRYNPEQTKVLSASQAF 180  
|  
DB 121 YFAOHGIIYVRSKAKFNEKEKRNKLOSVDAGTPMYLVIPEEGTRYNPEQTKVLSASQAF 180  
QY 181 AAORG 185  
|  
DB 181 AAORG 185  
RESULT 15  
US-09-338-907-135  
; Sequence 135, Application US/09338907  
; Patent No. 6265546  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Ilya, Chumakov  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: PROSTATE CANCER GENE  
; FILE REFERENCE: GENSET.18CPLCP  
; CURRENT APPLICATION NUMBER: US/09/338,907  
; EARLIER FILING DATE: 1999-06-23  
; EARLIER APPLICATION NUMBER: 08/996,306

; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: 60/099, 658  
; EARLIER FILING DATE: 1998-09-09  
; EARLIER APPLICATION NUMBER: 09/218, 207  
; EARLIER FILING DATE: 1998-12-22  
; NUMBER OF SEQ ID NOS: 578  
; SOFTWARE: Patent.pm  
; SEQ ID NO 135  
; LENGTH: 300  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: 81..83  
; OTHER INFORMATION: Box 1  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: 160..165  
; OTHER INFORMATION: Box 11  
US-09-338-907-135

Query Match 80.6%; Score 970; DB 4; Length 300;  
Best Local Similarity 100.0%; Pred. No. 3.1e-100;  
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRYLPSVVLIGTAPTYVLAWGVRLSAPFLPARYQALDDRLXCVQSMVLEFFENTG 60  
Db 1 MRYLPSVVLIGTAPTYVLAWGVRLSAPFLPARYQALDDRLXCVQSMVLEFFENTG 60  
OY 61 VOILLYGDLPRKNENIYLANHOSVDWIVADILAIRONALGHVRYVLEKGIKWLPLYGC 120  
Db 61 VOILLYGDLPRKNENIYLANHOSVDWIVADILAIRONALGHVRYVLEKGIKWLPLYGC 120  
OY 121 YFAOHGIIYVRSKAFENKRNKLOSTYVDAGTPLYVIFPEGTRYNPQOTKVLASQAF 180  
Db 121 YFAOHGIIYVRSKAFENKRNKLOSTYVDAGTPLYVIFPEGTRYNPQOTKVLASQAF 180  
OY 181 AAQR 184  
Db 181 AAQR 184

Search completed: August 28, 2002, 11:08:12  
Job time: 291 sec

**THIS PAGE BLANK (USPTO)**







14-SEP-2000; 2000US-0662191.  
PR 19-OCT-2000; 2000US-0693036.  
PR 29-NOV-2000; 2000US-0727344.  
XX  
PA (HUSE-) HUSEQ INC.  
XX  
PI Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX  
DR WPI: 2001-442253/47.  
DR N-PSDB: AA158747.  
XX  
PT Novel nucleic acids and polypeptides, useful for treating disorders  
PT such as central nervous system injuries -  
XX  
PS Example 4: SEQ ID NO 2736; 10078pp: English.  
XX  
XX The invention relates to human nucleic acids (AA157798-AA161369) and  
CC the encoded polypeptides (AA136642-AA142213) with nootropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/Inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.  
XX  
SQ Sequence 353 AA:

Query Match 100.0%; Score 1203; DB 22; Length 353;  
Best Local Similarity 100.0%; Pred. No. 4.6e-124;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRYLPSVVLGTAPRYVLAAGWRLLSAFLPARFQALDDRLXCYQSWLFFPENYTG 60  
DB 1 mryllpsvvlgtapryvlawgwrllsalfparfygalddrlycyqswvllffenytg 60  
OY 61 VOILLYGDLPKKNENITVLANHOSVDWIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120  
DB 61 vqilllygdlpkknenitvianhbstvcdwivadilairgnalghvryvllkeglkwlplygc 120  
OY 121 YFAOHGIGIVKRSKFNEMKRNKLOSVDAGTPMVLTFPESTRNPEQTKVLSAQAF 180  
DB 121 yfaohgigivkrsakfnemrnlklosyvdagtpmvltfpestrnpeqtkvlsasqaf 180  
OY 181 AAQRGIAVLKHLVTPRIKATHVAFDCMKNYLDAIVDTVTVVYVYESKDDGG 228  
DB 181 aaqrgiavlkhvltprikatvafdcmknyldaiydtvtvyvgyegkddgg 228

RESULT 4  
AAB85532  
ID AAB85532 strand: protein; 353 AA.  
AC AAB85532;  
XX  
XX 25-SEP-2001 (first entry)  
DE Human secreted protein (clone id HUTH084).  
XX  
XX Secreted protein; immunosuppressive; antiarthritic; antirheumatic;  
KW antiproliferative; cytostatic; cardiatic; vasotrophic; cerebroprotective;  
KW nootropic; neuroprotective; antibacterial; virocidic; fungicide; human;  
KW ophthalmological; gene therapy.

XX Homo sapiens.  
OS  
XX WO200155430-A1.  
PN  
XX  
PD 02-AUG-2001.  
PD  
XX  
PF 17-JAN-2001; 2001WO-US01431.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 12-SEP-2000; 2000US-0231968.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR, Olsen HS;  
PI Moore PA, Wei P, Edner R, Duan DR, Shi Y, Choi GH, Fiscella M;  
PI Ni J, Ruben SM, Barash SC;  
XX  
DR WPI: 2001-476220/51.  
DR N-PSDB: AAH46942.

17 isolated nucleic acid molecules encoding human secreted proteins,  
used to preventing, treating or ameliorating a medical condition -  
Claim 11: Page 450-451; 482pp; English.

The invention provides novel human secreted proteins and polynucleotides  
encoding them. The secreted proteins can be expressed by standard  
recombinant methodology. The secreted proteins and polynucleotides are  
used to prevent, treat or ameliorate a medical condition in e.g. humans,  
mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They can  
also be used in diagnosing a pathological condition. The antibodies to  
the proteins can also be used in alleviating symptoms associated with the  
disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme  
linked immunosorbent assays (ELISA). Disorders which are diagnosed or  
treated include autoimmune diseases e.g. rheumatoid arthritis,  
hyperproliferative disorders e.g. neoplasms of the breast or liver,  
cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
e.g. cerebral ischemia, angiodenesis, nervous system disorders e.g.  
Alzheimer's disease, infections caused by bacteria, viruses and fungi and  
ocular disorders e.g. corneal infection. The polypeptides can also be  
used to aid wound healing and epithelial cell proliferation, to prevent  
skin aging due to sunburn, to maintain organs before transplantation, for  
supporting cell culture of primary tissues, to regenerate tissues and in  
chemotaxis. The polypeptides can also be used as a food additive or  
preservative to increase or decrease storage capabilities. The present  
sequence represents a human secreted protein.

Query Match 100.0%; Score 1203; DB 22; Length 353;  
Best Local Similarity 100.0%; Pred. No. 4.6e-124;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRYLPSVVLGTAPRYVLAAGWRLLSAFLPARFQALDDRLXCYQSWLFFPENYTG 60  
DB 1 mryllpsvvlgtapryvlawgwrllsalfparfygalddrlycyqswvllffenytg 60  
OY 61 VOILLYGDLPKKNENITVLANHOSVDWIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120  
DB 61 vqilllygdlpkknenitvianhbstvcdwivadilairgnalghvryvllkeglkwlplygc 120  
OY 121 YFAOHGIGIVKRSKFNEMKRNKLOSVDAGTPMVLTFPESTRNPEQTKVLSAQAF 180  
DB 121 yfaohgigivkrsakfnemrnlklosyvdagtpmvltfpestrnpeqtkvlsasqaf 180  
OY 181 AAQRGIAVLKHLVTPRIKATHVAFDCMKNYLDAIVDTVTVVYVYESKDDGG 228  
DB 181 aaqrgiavlkhvltprikatvafdcmknyldaiydtvtvyvgyegkddgg 228

```

RESULT 5
AAB85552
ID AAB85552 standard; protein; 353 AA.
XX
AC AAB85552:
XX
DT 25-SEP-2001 (first entry)
XX
DE Human secreted protein (clone id HSLIA81).
XX
KW Secreted protein; immunosuppressive; antiarthritic; antirheumatic;
KW antiproliferative; cytosolic; cardiac; vasotropic; cerebroprotective;
KW neurotropic; neuroprotective; antibacterial; virucide; fungicide; human;
KW ophthalmological; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200155430-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01431.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 12-SEP-2000; 2000US-0231968.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Komatsu S, Baker KP, Birse CE, Soppet DR, Olsen HS;
PI Moore PA, Wei P, Edner R, Duan DR, Shi Y, Choi GH, Fischella M;
PI Ni J, Ruben SM, Barash SC;
XX
DR MPI; 2001-476220/51.
DR N-PSDB; AAH46962.
XX
PT 17 Isolated nucleic acid molecules encoding human secreted proteins,
PT used to preventing, treating or ameliorating a medical condition
XX
PS Claim 11; Page 465-466; 482pp; English.
XX
CC The invention provides novel human secreted proteins and polynucleotides
CC encoding them. The secreted proteins can be expressed by standard
CC recombinant methodology. The secreted proteins and polynucleotides are
CC used to prevent, treat or ameliorate a medical condition in e.g. humans,
CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They can
CC also be used in diagnosing a pathological condition. The antibodies to
CC the proteins can also be used in alleviating symptoms associated with the
CC disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme
CC linked immunosorbent assays (ELISA). Disorders which are diagnosed or
CC treated include autoimmune diseases e.g. Rheumatoid arthritis,
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
CC e.g. cerebral ischemia, angiogenesis, nervous system disorders e.g.
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and
CC ocular disorders e.g. corneal infection. The polypeptides can also be
CC used to aid wound healing and epithelial cell proliferation, to prevent
CC skin aging due to sunburn, to maintain organs before transplantation, for
CC supporting cell culture of primary tissues, to regenerate tissues and in
CC chemotaxis. The polypeptides can also be used as a food additive or
CC preservative to increase or decrease storage capabilities. The present
CC sequence represents a human secreted protein.
CC
XX
SQ Sequence 353 AA:

```

Query Match 100.0%; Score 1203; DB 22; Length 353;  
 Best Local Similarity 100.0%; Pred. No. 4.6e-124;  
 Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 MRLTSPSVLLGTAPTYVLAMGVRLSLAPLPAPFYQALDDRLRYCYQSVNVLFFEFNYTG 60
DB 1 MRLTSPSVLLGTAPTYVLAMGVRLSLAPLPAPFYQALDDRLRYCYQSVNVLFFEFNYTG 60

```

```

OY 61 VOILLXGDLPRKKNENIITYLANHSTVDMIVADILAIKRONALGHVRYLKEGLKMLPLXGC 120
DB 61 Vqlllygdldpknkenillylanhgstvdwivadilalrqnalghvryllkeglkmlplxgc 120
OY 121 YFAOHGCIYVRSKAKFKEMKRNKLSQSYVDAGTPMILVITPESTRNPEOTKXLSASQAR 180
DB 121 yfaohgciyvrskakfkemrnlksqsyvdagtpmilyvitpestrnpeotkxlsasqar 180
OY 181 AAGRLAVLKHVLTPIKATHVAFDCKKNVLDIAIVDTVYVEGKDDGG 228
DB 181 aagrlavlvkhvltprikatvafdcnknyldaiyvdtvvyegkddgg 228

```

```

RESULT 6
AAY36729
ID AAY36729 standard; protein; 364 AA.
XX
AC AAY36729;
XX
DT 27-SEP-1999 (first entry)
XX
DE Human PGI protein sequence.
XX
KW PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;
KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.
XX
OS Homo sapiens.
XX
PN WO932644-A2.
XX
PD 01-JUL-1999.
XX
PF 22-DEC-1998; 98WO-IB02133.
XX
PR 09-SEP-1998; 98US-0099658.
PR 22-DEC-1997; 97US-0396306.
XX
PA (GEST ) GENSET.
XX
PI Blumenfeld M, Bouguetelret L, Chumakov I, Cohen D;
XX
DR MPI; 1999-405178/34.
DR N-PSDB; AA200870.
XX
PT Use of a prostate cancer associated gene and biallelic markers
PT derived from it
XX
PS Claim 7; Page 190-191; 385pp; English.
XX
CC The invention relates to a mammalian PGI gene and protein, and a set of
CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are
CC used in a hybridisation assay, a sequencing assay, or in an
CC allele-specific amplification assay for determining the identity of a
CC nucleotide at a PGI-related biallelic marker. The methods can be used to
CC detect and to assess the risk of developing cancer or prostate cancer.
CC Early-stage diagnosis of prostate cancer relies on prostate specific
CC antigen (PSA) dosage. However, the effectiveness of this is limited due
CC to its inability to discriminate between malignant and non-malignant
CC affections of the organ. A need exists for both a reliable diagnostic
CC procedure which would enable early-stage diagnosis, and for preventative
CC and curative treatments of the disease. The PGI gene can be used for
CC detection of prostate cancer, and the risk of developing it in the
CC future, and can also be used to determine therapies for the disease.
CC
XX
SQ Sequence 364 AA:

```

Query Match 100.0%; Score 1203; DB 20; Length 364;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-124;  
 Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 MRLTSPSVLLGTAPTYVLAMGVRLSLAPLPAPFYQALDDRLRYCYQSVNVLFFEFNYTG 60

```

Db	12	mcylipsvlllgteapcylawgwrllsaafiparfygalddrllycvygsmliffenytg	71
Oy	61	VOIILIGDLPKKNENIITYLANHOSYVDMIVADIIAIONMAGHRYLKKEGKMLPLYGC	120
Db	72	vqlllygdllpxkenillylanhgstvdmivadialatqnaighrvylkegikwlpilygc	131
Oy	121	YFAOHGGIYKRSKAFENKEKEMNKLOSVDNCGPMYLVIPPEGKRYNPEQTKVLSAQAF	180
Db	132	yfaqgsgilyvkrtsakfnekemnrklsgsydagtcmlylvpfegrlrypedqkvlsasqaf	191
Oy	181	AAQGRLLAVLKHVLPRIKATIVAADCKNKNYDALIYDTPVYVEGDDSG	228
Db	192	aaqgylavlkhnvlprlkativaldcmknyldalydvtvyvegddsg	239

RESULT	7
AAAM50128	
ID	AAAM50128 standard; Protein; 364 AA
XX	
AC	AAAM50128;
DT	21-DEC-2001 (first entry)
XX	
DE	Human acyltransferase 27417.

KW Acyltransferase 27417; human; cell proliferation; cell migration;  
KW cell differentiation; ovary cancer; brain cancer; colon cancer;  
KW lung cancer; tumour; metastasis; sarcoma; carcinoma;  
KW adenocarcinoma; antitumour; lipid metabolism; diagnosis; therapy  
XX  
OS Homo sapiens.

FT	Key	Location/Qualifiers
FT	Peptide	1..49
FT		/label= Signal_peptide
FT	Protein	50..364
FT		/label= Mature_protein
FT	Domain	50..320
FT		/label= N-terminal_domain
FT	Domain	321..337
FT		/label= Transmembrane_domain
FT	Domain	338..364
FT		/label= C-terminal_domain
FT	Domain	71..363
FT		/label= Acyltransferase_domain
FT	Modified-site	68..71
FT		/note= "asn is N-glycosylated"
FT	Modified-site	241..244
FT		/note= "O-phosphorylated by cAMP- and CGMP-dependent protein kinase"
FT	Modified-site	11..13
FT		/note= "O-phosphorylated by protein kinase C"
FT	Modified-site	144..146
FT		/note= "O-phosphorylated by protein kinase C"
FT	Modified-site	205..207
FT		/note= "O-phosphorylated by protein kinase C"
FT	Modified-site	317..319
FT		/note= "O-phosphorylated by protein kinase C"
FT	Modified-site	361..363
FT		/note= "O-phosphorylated by protein kinase C"
FT	Modified-site	95..98
FT		/note= "O-phosphorylated by protein kinase C"
FT	Modified-site	158..161
FT		/note= "O-phosphorylated by casein kinase II"
FT	Modified-site	246..249
FT		/note= "O-phosphorylated by casein kinase II"
FT	Modified-site	82..89
FT		/note= "O-phosphorylated by tyrosine kinase"
FT	Modified-site	219..226
FT		/note= "O-phosphorylated by tyrosine kinase"
FT	Modified-site	23..28
FT		/note= "N-myristoylated"

FT	Modified-site	130..135
FT	/note="N-myristoylated"	
FT	Modified-site	330..335
FT	/note="N-myristoylated"	
FT	Modified-site	352..357
FT	/note="N-myristoylated"	
FT	Modified-site	230..343
FT	/note="Amidation"	
FT	Peptide	105..126
FT	/note="predicted leucine zipper"	

PN	W0200173051-A2.
XX	
PD	04-OCT-2001.

26-MAR-2001; 2001WO-US09633

PR 24-MAR-2000; 2000US-192092P

PA (MILL-) MILLENNIUM PHARM INC.

PI Meyers RA, Rudolph-Owen L, Macbeth KJ, Tsai F, Williamson M;

DR WPI; 2001-626267/72.

DR N-PSDB; AAH27094.

PT New polynucleotide encodes a polypeptide for the diagnosis and

PT polynucleotide encoding human acyltransferases designated 46743 and

PT 27417 -

PS Claim 9; Fig 8A-B; 150pp; English.

The present sequence is that of a novel human protein, termed 27417, which shows the structural characteristics of members of the acyltransferase family, including an acyltransferase domain. The invention provides novel acyltransferase 27417 nucleic acids and polypeptides, as well as methods for detecting their presence, and methods for screening for compounds that modulate their expression or activity. Such compounds can be used to treat conditions related to aberrant activity or expression of 27417 protein or nucleic acid, such as conditions involving deficient cellular proliferation, migration and/or differentiation. Also provided are methods for inhibiting the proliferation or migration, or inducing the killing, of a 27417-expressing cell, e.g. a hyperproliferative and/or metastatic cell, by contact with a compound that modulates the activity or expression of the 27417 protein or nucleic acid. The 27417-expressing cell is found in the healthy or diseased heart, blood vessels, kidney, skeletal muscle, brain or liver, or especially in a solid tumour, a soft tissue tumour or a metastatic lesion, a sarcoma, a carcinoma or an adenocarcinoma, and in particular in a hyperproliferative and/or metastatic cell found in ovarian, brain, colon or lung cancer. The compounds may also be used to modulate lipid metabolism in a 27417-expressing cell. Methods of disease diagnosis, e.g. by determining the presence of a genetic alteration in a 27417 polypeptide, and for evaluating the efficacy of a treatment of a disorder, are also provided.

AA	Sequence	364	AA:
SQ			

Query Match	100.0%	Score 1203;	DB 22;	Length 364;
Best Local Similarity	100.0%;	Pred. No. 4.8e-124;		
Matches 228; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

QY 1 MMYLLPSVLLGTAETVYLVANGVWELLASFLPAPFYALDDRLXCVYQSNVLEFFENYTG 60  
Db 12 MYLLPSVLLIGTAPCYLVAGVWRLLSAPFAPFYALDR LYCVYQSNVLLFFENYTG 71  
QY 61 VOILLGDLPRKNENI IYLANHQSIVDMIVADIIAIRONALGHNRYVLKEGLKMLPYGC 120  
Db 72 VVLLIGDGLPRKNENI IYLANHQSIVDMIVADIIAIFQNALGHNRYVLKEGLKMLPYGC 131

QY 121 YFAOHGSIYVKRSKAKENEKERNKLSQSYVDAGTPMWLVIEPEGRYRNEQTKVLSASQAF 180  
 |||  
 DB 132 Yfaqhggilyvkrsakfnekernklqsyvdagtpmylviefpegrlyrpeqtkvlsasqaf 191  
 QY 181 AAGRGALVLTGAPVLTPTKATHTVAFDCKKNLDATYDVTYVYEGKDDCG 228  
 |||  
 DB 192 aagrgalvltgavltprtkathvaIdcmknyldalvdlvlyegkddgg 239

## RESULT 8

AAU00665  
 ID AAU00665 standard; Protein: 364 AA.  
 XX  
 AC AAU00665;

DT 07-SEP-2001 (first entry)  
 XX

DE Human lysophosphatidic acid acyltransferase isoform LPAAT-epsilon.

XX Lysophosphatidic acid acyltransferase epsilon; LPAAT-epsilon; LPA; PA;  
 KM lysophosphatidic acid; phosphatidic acid; acylation; cellular activation;  
 KM phospholipid signaling; mitogenesis; inflammation; autoimmune disease;  
 KM oncology; cancer; obesity; gene therapy.

XX Homo sapiens.  
 OS  
 PN W0200134782-A1.

PD 17-MAY-2001.

PE 02-NOV-2000; 2000MO-US30193.

PR 09-NOV-1999; 99US-0436919.

PA (CELL-) CELL THERAPEUTICS INC.

PI Leung DW;

DR WPI: 2001-335920/35.

DR N-PSDB; AAS00649.

XX Novel isolated human isoform of lysophosphatidic acid  
 PT acyltransferase-epsilon useful for diagnostic, therapeutic and  
 PT screening purposes -  
 XX  
 PS Claim 1; Fig 1; 48pp; English.

XX The sequence represents a human lysophosphatidic acid acyltransferase  
 CC (LPAAT) isoform, LPAAT-epsilon. LPAAT catalyzes the acylation of  
 CC lysophosphatidic acid (LPA) to phosphatidic acid (PA). LPA and PA have  
 CC been identified as phospholipid signaling molecules that affect a wide  
 CC range of biological responses. PA is involved in cellular activation and  
 CC mitogenesis. Compounds that block PA generation and hence diminish lipid  
 CC biosynthesis and the signal involved in cell activation are of  
 CC therapeutic interest in the areas of inflammation and oncology (e.g.  
 CC autoimmune diseases and cancer) as well as obesity treatment.  
 CC LPAAT-epsilon and its corresponding DNA can be used in screening assays  
 CC to detect agents that stimulate or inhibit the activity of LPAAT and,  
 CC therefore, PA. The DNA is useful in tests to detect the presence or  
 CC expression of LPAAT-epsilon in relation to certain diseases and  
 CC conditions, and in disease prevention and treatment. The sequences of the  
 CC invention are also useful for diagnosis of diseases and conditions in  
 CC which the expression of LPAAT enzyme is abnormal.

XX Sequence 364 AA;

Query Match 100.0%; Score 1203; DB 22; Length 364;  
 Best local Similarity 100.0%; Pred. No. 4,8e-124;  
 Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYLLPSVVLGTAPVLTPTKATHTVAFDCKKNLDATYDVTYVYEGKDDCG 60

DB |||  
 12 myllpsvvlgtapvltgavltprtkathvaIdcmknyldalvdlvlyegkddgg 71  
 QY 61 VOILLYDDLPRKKNENITYLANHOSTVDWIVADIATIQNALGHVRYLKGKMLPYGCG 120  
 |||  
 DB 72 vqlllygdldpknkenillylanhgstvgwivadilaatgnaighvrylkgkmlpygcg 131  
 QY 121 YFAOHGSIYVKRSKAKENEKERNKLSQSYVDAGTPMWLVIEPEGRYRNEQTKVLSASQAF 180  
 |||  
 DB 132 Yfaqhggilyvkrsakfnekernklqsyvdagtpmylviefpegrlyrpeqtkvlsasqaf 191  
 QY 181 AAGRGALVLTGAPVLTPTKATHTVAFDCKKNLDATYDVTYVYEGKDDCG 228  
 |||  
 DB 192 aagrgalvltgavltprtkathvaIdcmknyldalvdlvlyegkddgg 239

## RESULT 9

AAB93595  
 ID AAB93595 standard; Protein: 353 AA.  
 XX

AC AAB93595;

DT 26-JUN-2001 (first entry)  
 XX

DE Human protein sequence SEQ ID NO:13028.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens.

OS  
 PN EP1074617-A2.

PD 07-FEB-2001.

PE 28-JUL-2000; 2000EP-0116126.

PR 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX WPI: 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs -  
 XX

PS Claim 8; SEQ ID 13028; 2537PP + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602  
 CC full-length cDNAs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and  
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length

CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human CDNA sequences; AAB92446 to  
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.

XX Sequence 353 AA:

Query Match 99.8%; Score 1200; DB 22; Length 353;  
 Best Local Similarity 99.6%; Pred. No. 9,8e-124;  
 Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRLLPSVLLGTAFTVYVLAAGVWRLLSAFLPARFYQALDDRLCYQSWVLEFFENYTG 60  
 Db 1 mryllpsvllgtaftvylawgwrrllsaflparfyqalddrlcyqswvleffenytg 60  
 OY 61 VQILLYGDLPKKNENIITYLANHSTVDWIVADILAIRONALGHVRYLKEGLKWLPLYGC 120  
 Db 61 vqillygdldpkknentiylanhgstvdwivadilairnalghvrylkeglkwlplygc 120  
 OY 121 YFAQHGCIYVKRSKAKFNEKEMRNKLOSVDAGTPMYLVIFPEGTRINPEQTKVLSASQAF 180  
 Db 121 yfaqhgciyvkrsakfnekemrnlqsyvdagtpmylvlfpegrtrynpeqtkvlsasqaf 180  
 OY 181 AAGRGIAVLKHYLTPRIKATHVAFDCKMKNYLDIYDVTVVYEGKDDGG 228  
 Db 181 aagrgiavlkhyltprikathvafdcmknyldaiydvttvvyegkddgg 228

RESULT 10  
 ID AAM41377 standard; Protein; 372 AA.

XX AAM41377;  
 AC AAM41377;  
 XX AAM41377;  
 DT 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 6308.  
 DE Human polypeptide SEQ ID NO 6308.

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia.

XX Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US34263.

XX 21-JAN-2000; 2000US-0488725.

XX 25-APR-2000; 2000US-0552317.

XX 09-JUL-2000; 2000US-0598042.

XX 19-JUL-2000; 2000US-0620312.

XX 03-AUG-2000; 2000US-0653450.

XX 14-SEP-2000; 2000US-0662191.

XX 19-OCT-2000; 2000US-0693036.

XX 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 PI Zhao QA, Zhou P, Goodrich R, Dzmanac RT;  
 XX WPI: 2001-442253/47.  
 DR N-PSDB; AAI60533.  
 XX

PT Novel nucleic acids and polypeptides, useful for treating disorders  
 PT such as central nervous system injuries -

XX Example 2; SEQ ID NO 6308; 10078bp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and  
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localized neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemia and  
 CC C.N.S disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.

XX Sequence 372 AA:

Query Match 99.1%; Score 1192; DB 22; Length 372;  
 Best Local Similarity 99.6%; Pred. No. 8.1e-123;  
 Matches 227; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 MRLLPSVLLGTAFTVYVLAAGVWRLLSAFLPARFYQALDDRLCYQSWVLEFFENYTG 60  
 Db 20 mryllpsvllgtaftvylawgwrrllsaflparfyqalddrlcyqswvleffenytg 79  
 OY 61 VQILLYGDLPKKNENIITYLANHSTVDWIVADILAIRONALGHVRYLKEGLKWLPLYGC 120  
 Db 80 vqillygdldpkknentiylanhgstvdwivadilairnalghvrylkeglkwlplygw 139  
 OY 121 YFAQHGCIYVKRSKAKFNEKEMRNKLOSVDAGTPMYLVIFPEGTRINPEQTKVLSASQAF 180  
 Db 140 yfaqhgciyvkrsakfnekemrnlqsyvdagtpmylvlfpegrtrynpeqtkvlsasqaf 199  
 OY 181 AAGRGIAVLKHYLTPRIKATHVAFDCKMKNYLDIYDVTVVYEGKDDGG 228  
 Db 200 aagrgiavlkhyltprikathvafdcmknyldaiydvttvvyegkddgg 247

RESULT 11

XX AAY36741  
 ID AAY36741 standard; Protein; 354 AA.

XX AAY36741;

XX 27-SEP-1999 (first entry)

XX Mouse PGI protein sequence.

XX PGI gene; diallelic marker; mouse; PSA; PGI-related diallelic marker;  
 KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.

XX Mus musculus.

XX WO9932644-A2.

XX 01-JUL-1999.

XX 22-DEC-1998; 98WO-IB02133.

XX 09-SEP-1998; 98US-0099658.

XX 22-DEC-1997; 97US-0996306.

XX (GEST ) GENSET.

XX Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;  
 PI

XX WPI: 1999-405178/34.  
 DR N-PSDB: AAZ00928, AAZ01027.  
 XX  
 PT Use of a prostate cancer associated gene and biallelic markers  
 derived from it  
 PS Claim 7, Page 215, 385pp; English.  
 XX  
 CC The invention relates to a mammalian pgl gene and protein, and a set of  
 CC biallelic markers. The pgl polynucleotide and biallelic markers are  
 CC used in a hybridisation assay, a sequencing assay, or in an  
 CC allele-specific amplification assay for determining the identity of a  
 CC nucleotide at a pgl-related biallelic marker. The methods can be used to  
 CC detect and to assess the risk of developing cancer or prostate cancer.  
 CC Early-stage diagnosis of prostate cancer relies on prostate specific  
 CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
 CC to its inability to discriminate between malignant and non-malignant  
 CC affections of the organ. A need exists for both a reliable diagnostic  
 CC procedure which would enable early-stage diagnosis, and for preventative  
 CC and curative treatments of the disease. The pgl gene can be used for  
 CC detection of prostate cancer, and the risk of developing it in the  
 CC future, and can also be used to determine therapies for the disease.  
 XX  
 SQ Sequence 354 AA;

Query Match 86.5%; Score 1040; DB 20; Length 354;  
 Best Local Similarity 85.9%; Pred. No. 4.4e-106;  
 Matches 195; Conservative 19; Mismatches 13; Indels 0; Gaps 0;  
 QY 1 MRYLLPSVLLGTAPTYVLANGVWRLLSAFLPARFYQALDRLCYVQSMVLPFFENYTG 60  
 Db 1 mryllpsvlllgtaptyvllangvwrllsafiparfyaldrlycvygsmlfffenytg 60  
 QY 61 VOILLYGDLPKNKENIYILANHOSTVDWIVADILAIRONALGHWRYVLKEGLKMLPLYGC 120  
 Db 61 vqilllygdlpknkenillylanhostvdwivadilaaironalghwryvlkeglkmlplygc 120  
 QY 121 YFAOHGCIYVRSKAFKFNKEKRNKLQSYVDAGTPMYLVIPEFGTRYNPEQTKVLSASQAF 180  
 Db 121 yfaohgciyvrsakfknkernklqsyvvdagtpmylvifepgtrynpegtkvlsasqaf 180  
 QY 181 AAOGLAVLKHVLPRIKATVAFDCMKKNYDAIYDVTVYEGKDDG 227  
 Db 181 aagrlavlkhlvprlkatvafdcmkknnydaiydvttvyegkddg 227

RESULT 12  
 AAY36753  
 ID AAY36753 standard; Protein: 185 AA.  
 XX  
 AC AAY36753;  
 XX  
 DT 27-SEP-1999 (first entry)  
 XX  
 DE pgl splice variant allele protein sequence.  
 XX  
 KW pgl gene; biallelic marker; human; PSA; pgl-related biallelic marker;  
 KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9932644-A2.  
 XX  
 PD 01-JUL-1999.  
 XX  
 PF 22-DEC-1998; 98WO-1B02133.  
 XX  
 PR 09-SEP-1998; 98US-0099658.  
 PR 22-DEC-1997; 97US-0996306.  
 XX  
 PA (GEST ) GENSET.

XX Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;  
 PI WPI: 1999-405178/34.  
 XX DR N-PSDB: AAZ00979.  
 XX  
 PT Use of a prostate cancer associated gene and biallelic markers  
 derived from it  
 PS Claim 7, Page 255, 385pp; English.  
 XX  
 CC The invention relates to a mammalian pgl gene and protein, and a set of  
 CC biallelic markers. The pgl polynucleotide and biallelic markers are  
 CC used in a hybridisation assay, a sequencing assay, or in an  
 CC allele-specific amplification assay for determining the identity of a  
 CC nucleotide at a pgl-related biallelic marker. The methods can be used to  
 CC detect and to assess the risk of developing cancer or prostate cancer.  
 CC Early-stage diagnosis of prostate cancer relies on prostate specific  
 CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
 CC to its inability to discriminate between malignant and non-malignant  
 CC affections of the organ. A need exists for both a reliable diagnostic  
 CC procedure which would enable early-stage diagnosis, and for preventative  
 CC and curative treatments of the disease. The pgl gene can be used for  
 CC detection of prostate cancer, and the risk of developing it in the  
 CC future, and can also be used to determine therapies for the disease.  
 XX  
 SQ Sequence 185 AA;

Query Match 81.1%; Score 976; DB 20; Length 185;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-99;  
 Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MRYLLPSVLLGTAPTYVLANGVWRLLSAFLPARFYQALDRLCYVQSMVLPFFENYTG 60  
 Db 1 mryllpsvlllgtaptyvllangvwrllsafiparfyaldrlycvygsmlfffenytg 60  
 QY 61 VOILLYGDLPKNKENIYILANHOSTVDWIVADILAIRONALGHWRYVLKEGLKMLPLYGC 120  
 Db 61 vqilllygdlpknkenillylanhostvdwivadilaaironalghwryvlkeglkmlplygc 120  
 QY 121 YFAOHGCIYVRSKAFKFNKEKRNKLQSYVDAGTPMYLVIPEFGTRYNPEQTKVLSASQAF 180  
 Db 121 yfaohgciyvrsakfknkernklqsyvvdagtpmylvifepgtrynpegtkvlsasqaf 180  
 QY 181 AAOGRG 185  
 Db 181 aagrg 185

RESULT 13  
 AAY36751  
 ID AAY36751 standard; Protein: 315 AA.  
 XX  
 AC AAY36751;  
 XX  
 DT 27-SEP-1999 (first entry)  
 XX  
 DE pgl splice variant allele protein sequence.  
 XX  
 KW pgl gene; biallelic marker; human; PSA; pgl-related biallelic marker;  
 KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9932644-A2.  
 XX  
 PD 01-JUL-1999.  
 XX  
 PF 22-DEC-1998; 98WO-1B02133.  
 XX  
 PR 09-SEP-1998; 98US-0099658.  
 PR 22-DEC-1997; 97US-0996306.



XX (GEST ) GENSET.  
PA Blumentfeld M, Bougueleret L, Chumakov I, Cohen D;  
XX WPI: 1999-405178/34.  
DR N-PSDB: AA200977.  
XX  
PT Use of a prostate cancer associated gene and biallelic markers  
XX derived from it  
PS Claim 7; Page 253-254; 385pp; English.  
XX  
CC The invention relates to a mammalian PGI gene and protein, and a set of  
CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
CC used in a hybridisation assay, a sequencing assay, or in an  
CC allele-specific amplification assay for determining the identity of a  
CC nucleotide at a PGI-related biallelic marker. The methods can be used to  
CC detect and to assess the risk of developing cancer or prostate cancer.  
CC Early-stage diagnosis of prostate cancer relies on prostate specific  
CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
CC to its inability to discriminate between malignant and non-malignant  
CC affections of the organ. A need exists for both a reliable diagnostic  
CC procedure which would enable early-stage diagnosis, and for preventative  
CC and curative treatments of the disease. The PGI gene can be used for  
CC detection of prostate cancer, and the risk of developing it in the  
CC future, and can also be used to determine therapies for the disease.  
XX  
SQ Sequence 315 AA;

Query Match 81.1%; Score 976; DB 20; Length 315;  
Best Local Similarity 100.0%; Pred. No. 4.3e-99;  
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRYLLPSVVLGTAPTYVLAMGWRLISAFLPARYQALDDRLCYQSVLFFENYTG 60  
DB 1 mryllpsvvlgtaptyvlamgwrlisaflparyqalddrlcyqsvmlffenytg 60  
OY 61 VQILLYGDLPRKKNENTIIYLANHOSTVDWIVADILAIROMALGHVRYLKEGLKWLPLYGC 120  
DB 61 vqillygdlprkknentiiylanhostvdwivadilairgnalghvrylkeglkwlplygc 120  
OY 121 YFAOHGCIYKRSKAKFNKEMKRLKLSYVDAGTPMVLVIFPEGTRNPEOTKVLASQAF 180  
DB 121 yfaohgciykrskakfnkemrnlklsyvdagtpmvlvivfpegtrnpeotkvlasqaf 180  
OY 181 AAQRG 185  
DB 181 aagrg 185

## RESULT 14

AA36752  
ID AAY36752 standard; Protein: 300 AA.

AC AAY36752;

DT 27-SEP-1999 (first entry)

DE PGI splice variant allele protein sequence.

XX PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;  
KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.

OS Homo sapiens.

PN WO932644-A2.

PD 01-JUL-1999.

PF 22-DEC-1998; 98WO-IB02133.

XX

PR 09-SEP-1998; 98US-0099658.  
PR 22-DEC-1997; 97US-0996306.  
XX

PA (GEST ) GENSET.

PI Blumentfeld M, Bougueleret L, Chumakov I, Cohen D;

XX WPI: 1999-405178/34.

DR N-PSDB: AA200978.

XX Use of a prostate cancer associated gene and biallelic markers  
XX derived from it  
XX

PS Claim 7; Page 254-255; 385pp; English.

XX  
CC The invention relates to a mammalian PGI gene and protein, and a set of  
CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
CC used in a hybridisation assay, a sequencing assay, or in an  
CC allele-specific amplification assay for determining the identity of a  
CC nucleotide at a PGI-related biallelic marker. The methods can be used to  
CC detect and to assess the risk of developing cancer or prostate cancer.  
CC Early-stage diagnosis of prostate cancer relies on prostate specific  
CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
CC to its inability to discriminate between malignant and non-malignant  
CC affections of the organ. A need exists for both a reliable diagnostic  
CC procedure which would enable early-stage diagnosis, and for preventative  
CC and curative treatments of the disease. The PGI gene can be used for  
CC detection of prostate cancer, and the risk of developing it in the  
CC future, and can also be used to determine therapies for the disease.  
XX  
SQ Sequence 300 AA;

Query Match 80.6%; Score 970; DB 20; Length 300;  
Best Local Similarity 100.0%; Pred. No. 1.8e-98;

Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRYLLPSVVLGTAPTYVLAMGWRLISAFLPARYQALDDRLCYQSVLFFENYTG 60  
DB 1 mryllpsvvlgtaptyvlamgwrlisaflparyqalddrlcyqsvmlffenytg 60  
OY 61 VQILLYGDLPRKKNENTIIYLANHOSTVDWIVADILAIROMALGHVRYLKEGLKWLPLYGC 120  
DB 61 vqillygdlprkknentiiylanhostvdwivadilairgnalghvrylkeglkwlplygc 120  
OY 121 YFAOHGCIYKRSKAKFNKEMKRLKLSYVDAGTPMVLVIFPEGTRNPEOTKVLASQAF 180  
DB 121 yfaohgciykrskakfnkemrnlklsyvdagtpmvlvivfpegtrnpeotkvlasqaf 180  
OY 181 AAQR 184  
DB 181 aagr 184

## RESULT 15

AA36744  
ID AAY36744 standard; Protein: 291 AA.

AC AAY36744;

DT 27-SEP-1999 (first entry)

DE PGI splice variant allele protein sequence.

XX PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;  
KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.

OS Homo sapiens.

PN WO932644-A2.

PD 01-JUL-1999.

XX



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:14:11 ; Search time 306.96 Seconds

(Without alignments)  
261.440 Million cell updates/sec

Title: US-09-853-526-70

Perfect score: 1203  
Sequence: 1 MRYLPSVVLGTAPIYVLA.....NYLDATYVYEGKDDG 228

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 3502263 segs, 351980561 residues

Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents\_AA\_Main:\*

- 1: /cgn2\_6/ptodata/2/paa/PCRUS\_COMB.pep:\*
- 2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep:\*
- 3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep:\*
- 4: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep:\*
- 5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep:\*
- 6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep:\*
- 7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep:\*
- 8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep:\*
- 9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep:\*
- 10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep:\*
- 11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep:\*
- 12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep:\*
- 13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep:\*
- 14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep:\*
- 15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep:\*
- 16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep:\*
- 17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep:\*
- 18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep:\*
- 19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep:\*
- 20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep:\*
- 21: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep:\*
- 22: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep:\*
- 23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep:\*
- 24: /cgn2\_6/ptodata/2/paa/US100\_COMB.pep:\*
- 25: /cgn2\_6/ptodata/2/paa/US101\_COMB.pep:\*
- 26: /cgn2\_6/ptodata/2/paa/US60\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1203	100.0	228	US-09-853-526-70	Sequence 70, Appl
2	1203	100.0	228	US-09-901-484A-70	Sequence 70, Appl
3	1203	100.0	228	US-60-099-658-70	Sequence 70, Appl
4	1203	100.0	353	PCT-US01-01431-59	Sequence 59, Appl
5	1203	100.0	353	PCT-US01-01431-79	Sequence 79, Appl
6	1203	100.0	353	PCT-US01-11988-1474	Sequence 1474, Ap
7	1203	100.0	353	PCT-US01-11988-1475	Sequence 1475, Ap

8	1203	100.0	353	18	US-09-488-725A-2736	Sequence 2736, Ap
9	1203	100.0	353	22	US-09-833-245-1474	Sequence 1474, Ap
10	1203	100.0	353	22	US-09-833-245-1475	Sequence 1475, Ap
11	1203	100.0	353	22	US-09-853-526-4	Sequence 4, Appl1
12	1203	100.0	353	22	US-09-901-484A-4	Sequence 59, Appl
13	1203	100.0	353	23	US-09-915-582-59	Sequence 79, Appl
14	1203	100.0	353	23	US-09-915-582-79	Sequence 4, Appl1
15	1203	100.0	353	26	US-60-099-658-4	Sequence 7, Appl1
16	1203	100.0	364	18	US-09-436-919-1	Sequence 5, Appl1
17	1203	100.0	364	22	US-09-817-910-7	Sequence 5, Appl1
18	1203	100.0	364	22	US-09-853-526-5	Sequence 5, Appl1
19	1203	100.0	364	23	US-09-901-484A-5	Sequence 5, Appl1
20	1203	100.0	364	26	US-60-099-658-5	Sequence 6308, Ap
21	1192	99.1	372	18	US-09-488-725A-6308	Sequence 110, App
22	1066.5	88.7	450	26	US-60-245-722-110	Sequence 309, App
23	1043.5	86.7	1032	26	US-60-212-413-309	Sequence 248, App
24	1043.5	86.7	1032	26	US-60-229-518-248	Sequence 74, Appl
25	1040	86.5	354	22	US-09-853-526-74	Sequence 74, Appl
26	1040	86.5	354	23	US-09-901-484A-74	Sequence 74, Appl
27	1040	86.5	354	26	US-60-099-658-74	Sequence 136, App
28	976	81.1	185	22	US-09-853-526-136	Sequence 136, App
29	976	81.1	185	23	US-09-901-484A-136	Sequence 134, App
30	976	81.1	315	22	US-09-901-484A-134	Sequence 135, App
31	976	81.1	315	23	US-09-853-526-135	Sequence 135, App
32	970	80.6	300	22	US-09-901-484A-135	Sequence 127, App
33	970	80.6	300	23	US-09-853-526-127	Sequence 127, App
34	830	69.0	291	22	US-09-901-484A-127	Sequence 133, App
35	830	69.0	291	23	US-09-853-526-133	Sequence 133, App
36	820	68.2	182	22	US-09-901-484A-133	Sequence 128, App
37	820	68.2	182	23	PCT-US01-01327-47	Sequence 128, App
38	759	63.1	269	1	US-09-853-526-128	Sequence 126, App
39	657	54.6	261	23	US-09-901-484A-128	Sequence 126, App
40	657	54.6	261	22	US-09-853-526-126	Sequence 126, App
41	597	49.6	238	22	US-09-901-484A-126	Sequence 5, Appl1
42	597	49.6	238	23	US-09-436-919-5	Sequence 132, App
43	594	49.4	176	18	US-09-853-526-132	Sequence 132, App
44	449	37.3	97	22		
45	449	37.3	97	23		

#### ALIGNMENTS

RESULT 1  
US-09-853-526-70  
; Sequence 70, Application US/09853526  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Ilyia, Chumakov  
; APPLICANT: Bouguetel, Lydie  
; TITLE OF INVENTION: PROSTATE CANCER GENE  
; FILE REFERENCE: GENSET.18CP1CP  
; CURRENT APPLICATION NUMBER: US/09/853,526  
; CURRENT FILING DATE: 2001-05-11  
; PRIOR APPLICATION NUMBER: 09/338,907  
; PRIOR FILING DATE: 1999-06-23  
; PRIOR APPLICATION NUMBER: 08/996,306  
; PRIOR FILING DATE: 1997-12-22  
; PRIOR APPLICATION NUMBER: 60/099,658  
; PRIOR FILING DATE: 1998-09-09  
; PRIOR APPLICATION NUMBER: 09/218,207  
; PRIOR FILING DATE: 1998-12-22  
; NUMBER OF SEQ ID NOS: 578  
; SOFTWARE: Patent.pm  
; SEQ ID NO 70  
; LENGTH: 228  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-853-526-70  
Query Match 100.0%; Score 1203; DB 22; Length 228;

Best Local Similarity 100.0%; Pred. No. 7.7e-120;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MRYLLPSVLLGTAPTYVLAAGWVRLSAFLPARFYQALDRLCYVQSMVLEFFENYTG 60
    |||||||
DB 1 MRYLLPSVLLGTAPTYVLAAGWVRLSAFLPARFYQALDRLCYVQSMVLEFFENYTG 60
QY 61 VOILLYGDLPKKNENIYLANHOSYVDMIVADILAIRONALGHVRYVLEKGLKMLPLXGC 120
    |||||||
DB 61 VOILLYGDLPKKNENIYLANHOSYVDMIVADILAIRONALGHVRYVLEKGLKMLPLXGC 120
QY 121 YFAOHGIIYVKSARFENEKEMRNKLOSVDAGTPLYVIFPEGTRYNPEQTKVLSASQAF 180
    |||||||
DB 121 YFAOHGIIYVKSARFENEKEMRNKLOSVDAGTPLYVIFPEGTRYNPEQTKVLSASQAF 180
QY 181 AAOGLAVLKHVLTPIKATHVAFDCMKNYLDIYDVTVYVEGKDDGG 228
    |||||||
DB 181 AAOGLAVLKHVLTPIKATHVAFDCMKNYLDIYDVTVYVEGKDDGG 228
```

## RESULT 2

US-09-901-484A-70  
; Sequence 70, Application US/09901484A  
; GENERAL INFORMATION:

APPLICANT: Cohen, Daniel  
APPLICANT: Blumenfeld, Marta  
APPLICANT: Chumakov, Ilya  
APPLICANT: Bougueleret, Lydie  
TITLE OF INVENTION: Prostate Cancer Gene  
FILE REFERENCE: GEN-T11XC3D2  
CURRENT APPLICATION NUMBER: US/09/901,484A  
CURRENT FILING DATE: 2001-07-09  
PRIOR APPLICATION NUMBER: US 08/996,306  
PRIOR FILING DATE: 1997-12-22  
PRIOR APPLICATION NUMBER: US 60/099,658  
PRIOR FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: US 09/218,207  
PRIOR FILING DATE: 1998-12-22  
PRIOR APPLICATION NUMBER: US 09/338,907  
PRIOR FILING DATE: 1999-06-23  
PRIOR APPLICATION NUMBER: US 09/853,526  
PRIOR FILING DATE: 2001-05-11  
NUMBER OF SEQ ID NOS: 578  
SOFTWARE: patentln version 3.1  
SEQ ID NO 70  
LENGTH: 228  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (5208)..(5208)  
OTHER INFORMATION: n = a, c, g, or t.

US-09-901-484A-70

Query Match 100.0%; Score 1203; DB 23; Length 228;  
Best Local Similarity 100.0%; Pred. No. 7.7e-120;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MRYLLPSVLLGTAPTYVLAAGWVRLSAFLPARFYQALDRLCYVQSMVLEFFENYTG 60
    |||||||
DB 1 MRYLLPSVLLGTAPTYVLAAGWVRLSAFLPARFYQALDRLCYVQSMVLEFFENYTG 60
QY 61 VOILLYGDLPKKNENIYLANHOSYVDMIVADILAIRONALGHVRYVLEKGLKMLPLXGC 120
    |||||||
DB 61 VOILLYGDLPKKNENIYLANHOSYVDMIVADILAIRONALGHVRYVLEKGLKMLPLXGC 120
QY 121 YFAOHGIIYVKSARFENEKEMRNKLOSVDAGTPLYVIFPEGTRYNPEQTKVLSASQAF 180
    |||||||
DB 121 YFAOHGIIYVKSARFENEKEMRNKLOSVDAGTPLYVIFPEGTRYNPEQTKVLSASQAF 180
QY 181 AAOGLAVLKHVLTPIKATHVAFDCMKNYLDIYDVTVYVEGKDDGG 228
    |||||||
```

DB 181 AAOGLAVLKHVLTPIKATHVAFDCMKNYLDIYDVTVYVEGKDDGG 228

## RESULT 3

US-60-099-658-70  
; Sequence 70, Application US/60099658  
; GENERAL INFORMATION:

APPLICANT: Cohen, Daniel  
APPLICANT: Chumakov, Ilya  
APPLICANT: Blumenfeld, Marta  
APPLICANT: Bougueleret, Lydie  
TITLE OF INVENTION: Prostate cancer gene  
NUMBER OF SEQUENCES: 99  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobbe, Martens, Olson & Bear  
STREET: 501 West Broadway  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92101-3505

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Win95  
SOFTWARE: Word

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/60/099,658  
FILING DATE:

CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:

NAME: Israelson, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: GENSET.018APR  
TELEPHONE: (619) 235-8550  
TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 70:  
SEQUENCE CHARACTERISTICS:

LENGTH: 228 amino acids  
TYPE: AMINO ACID

STRANDEDNESS: SINGLE  
TOPOLOGY: LINEAR

MOLECULE TYPE: Protein  
ORIGINAL SOURCE:

ORGANISM: Homo sapiens  
US-60-099-658-70

Query Match 100.0%; Score 1203; DB 26; Length 228;  
Best Local Similarity 100.0%; Pred. No. 7.7e-120;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MRYLLPSVLLGTAPTYVLAAGWVRLSAFLPARFYQALDRLCYVQSMVLEFFENYTG 60
    |||||||
DB 1 MRYLLPSVLLGTAPTYVLAAGWVRLSAFLPARFYQALDRLCYVQSMVLEFFENYTG 60
QY 61 VOILLYGDLPKKNENIYLANHOSYVDMIVADILAIRONALGHVRYVLEKGLKMLPLXGC 120
    |||||||
DB 61 VOILLYGDLPKKNENIYLANHOSYVDMIVADILAIRONALGHVRYVLEKGLKMLPLXGC 120
QY 121 YFAOHGIIYVKSARFENEKEMRNKLOSVDAGTPLYVIFPEGTRYNPEQTKVLSASQAF 180
    |||||||
DB 121 YFAOHGIIYVKSARFENEKEMRNKLOSVDAGTPLYVIFPEGTRYNPEQTKVLSASQAF 180
QY 181 AAOGLAVLKHVLTPIKATHVAFDCMKNYLDIYDVTVYVEGKDDGG 228
    |||||||
DB 181 AAOGLAVLKHVLTPIKATHVAFDCMKNYLDIYDVTVYVEGKDDGG 228
```

## RESULT 4

PCT-US01-01431-59  
; Sequence 59, Application PC/TUS0101431  
; GENERAL INFORMATION:

```
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: 17 human secreted proteins
; FILE REFERENCE: PS723PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01431
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/231,968
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-01431-59
```

```
Query Match          100.0%; Score 1203; DB 1: Length 353;
Best Local Similarity 100.0%; Pred. No. 1.4e-119;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 MRYLPSVVLGTAFTYVLAAGWRLLSAFLPARFYQALDDRLCYCYQSMTLFFPENYTG 60
    |||||||
DB 1 MRYLPSVVLGTAFTYVLAAGWRLLSAFLPARFYQALDDRLCYCYQSMTLFFPENYTG 60
OY 61 VOILLYGDLPKKNENIYYLANHSTVDWIVADILAIRQNALGHVRYLKEGLKWLPLYGC 120
    |||||||
DB 61 VOILLYGDLPKKNENIYYLANHSTVDWIVADILAIRQNALGHVRYLKEGLKWLPLYGC 120
OY 121 YFAOHGGIYVKSAREKEMRNKLOSVDAGTPMYLVIFPESTRNPEOTKVLASQAF 180
    |||||||
DB 121 YFAOHGGIYVKSAREKEMRNKLOSVDAGTPMYLVIFPESTRNPEOTKVLASQAF 180
OY 121 YFAOHGGIYVKSAREKEMRNKLOSVDAGTPMYLVIFPESTRNPEOTKVLASQAF 180
    |||||||
DB 121 YFAOHGGIYVKSAREKEMRNKLOSVDAGTPMYLVIFPESTRNPEOTKVLASQAF 180
OY 181 AAORGLAVLKHVLTPIKATHVAFDCMKNYLDIYDTVYVEGKDDG 228
    |||||||
DB 181 AAORGLAVLKHVLTPIKATHVAFDCMKNYLDIYDTVYVEGKDDG 228
```

```
RESULT 5
PCT-US01-01431-79
; Sequence 79, Application PC/TUS0101431
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: 17 human secreted proteins
; FILE REFERENCE: PS723PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01431
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/231,968
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 79
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-01431-79
```

```
Query Match          100.0%; Score 1203; DB 1: Length 353;
Best Local Similarity 100.0%; Pred. No. 1.4e-119;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 MRYLPSVVLGTAFTYVLAAGWRLLSAFLPARFYQALDDRLCYCYQSMTLFFPENYTG 60
    |||||||
DB 1 MRYLPSVVLGTAFTYVLAAGWRLLSAFLPARFYQALDDRLCYCYQSMTLFFPENYTG 60
```

```
OY 61 VOILLYGDLPKKNENIYYLANHSTVDWIVADILAIRQNALGHVRYLKEGLKWLPLYGC 120
    |||||||
DB 61 VOILLYGDLPKKNENIYYLANHSTVDWIVADILAIRQNALGHVRYLKEGLKWLPLYGC 120
OY 121 YFAOHGGIYVKSAREKEMRNKLOSVDAGTPMYLVIFPESTRNPEOTKVLASQAF 180
    |||||||
DB 121 YFAOHGGIYVKSAREKEMRNKLOSVDAGTPMYLVIFPESTRNPEOTKVLASQAF 180
OY 181 AAORGLAVLKHVLTPIKATHVAFDCMKNYLDIYDTVYVEGKDDG 228
    |||||||
DB 181 AAORGLAVLKHVLTPIKATHVAFDCMKNYLDIYDTVYVEGKDDG 228
```

```
RESULT 6
PCT-US01-11988-1474
; Sequence 1474, Application PC/TUS0111988
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: PCT/US01/11988
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1474
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-11988-1474
```

```
Query Match          100.0%; Score 1203; DB 1: Length 353;
Best Local Similarity 100.0%; Pred. No. 1.4e-119;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 MRYLPSVVLGTAFTYVLAAGWRLLSAFLPARFYQALDDRLCYCYQSMTLFFPENYTG 60
    |||||||
DB 1 MRYLPSVVLGTAFTYVLAAGWRLLSAFLPARFYQALDDRLCYCYQSMTLFFPENYTG 60
OY 61 VOILLYGDLPKKNENIYYLANHSTVDWIVADILAIRQNALGHVRYLKEGLKWLPLYGC 120
    |||||||
DB 61 VOILLYGDLPKKNENIYYLANHSTVDWIVADILAIRQNALGHVRYLKEGLKWLPLYGC 120
OY 121 YFAOHGGIYVKSAREKEMRNKLOSVDAGTPMYLVIFPESTRNPEOTKVLASQAF 180
    |||||||
DB 121 YFAOHGGIYVKSAREKEMRNKLOSVDAGTPMYLVIFPESTRNPEOTKVLASQAF 180
OY 121 YFAOHGGIYVKSAREKEMRNKLOSVDAGTPMYLVIFPESTRNPEOTKVLASQAF 180
    |||||||
DB 121 YFAOHGGIYVKSAREKEMRNKLOSVDAGTPMYLVIFPESTRNPEOTKVLASQAF 180
OY 181 AAORGLAVLKHVLTPIKATHVAFDCMKNYLDIYDTVYVEGKDDG 228
    |||||||
DB 181 AAORGLAVLKHVLTPIKATHVAFDCMKNYLDIYDTVYVEGKDDG 228
```

```
RESULT 7
PCT-US01-11988-1475
; Sequence 1475, Application PC/TUS0111988
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: PCT/US01/11988
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
```

NUMBER OF SEQ ID NOS: 2267  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO: 1475  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
PCT-US01-11986-1475

Query Match 100.0%; Score 1203; DB 1; Length 353;  
Best Local Similarity 100.0%; Pred. No. 1.4e-119;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTAFTYVLANGVWRLSAPLPAFYQALDRRLCYQSMVLFFENYTG 60  
DB 1 MRYLLPSVLLGTAFTYVLANGVWRLSAPLPAFYQALDRRLCYQSMVLFFENYTG 60  
QY 61 VOILLYGDLPKKNENIYLANHSTVDMIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120  
DB 61 VOILLYGDLPKKNENIYLANHSTVDMIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120  
QY 121 YFAOHGCIYVRSKAFNEKEKRNKLSQSYVDAGTMYLVIPEEGTRYNPEQTKVLSASQAF 180  
DB 121 YFAOHGCIYVRSKAFNEKEKRNKLSQSYVDAGTMYLVIPEEGTRYNPEQTKVLSASQAF 180  
QY 181 AAQGLAVLKHVLTPIKATHVAFDCKMKNYLDATYDVTVYEGKDDG 228  
DB 181 AAQGLAVLKHVLTPIKATHVAFDCKMKNYLDATYDVTVYEGKDDG 228

RESULT 8  
US-09-488-725A-2736  
Sequence 2736, Application US/09488725A  
GENERAL INFORMATION:  
APPLICANT: Hyseq Inc  
TITLE OF INVENTION: Novel Nucleic Acid and Polypeptides  
FILE REFERENCE: 784PLPCT  
CURRENT APPLICATION NUMBER: US/09/488, 725A  
CURRENT FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: US/09/488, 725  
PRIOR FILING DATE: 2000-01-21  
PRIOR APPLICATION NUMBER: US09/552, 317  
PRIOR FILING DATE: 2000-04-25  
PRIOR APPLICATION NUMBER: US09/598, 042  
PRIOR FILING DATE: 2000-06-20  
PRIOR APPLICATION NUMBER: US09/620, 312  
PRIOR FILING DATE: 2000-07-19  
PRIOR APPLICATION NUMBER: US09/653, 450  
PRIOR FILING DATE: 2000-08-31  
PRIOR APPLICATION NUMBER: US09/662, 191  
PRIOR FILING DATE: 2000-09-14  
PRIOR APPLICATION NUMBER: US09/693, 036  
PRIOR FILING DATE: 2000-10-19  
PRIOR APPLICATION NUMBER: US09/727, 344  
PRIOR FILING DATE: 2000-11-29  
NUMBER OF SEQ ID NOS: 7144  
SOFTWARE: PL\_FL\_genes\_b Versions 1.0  
SEQ ID NO 2736  
LENGTH: 354  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-488-725A-2736

Query Match 100.0%; Score 1203; DB 18; Length 353;  
Best Local Similarity 100.0%; Pred. No. 1.4e-119;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTAFTYVLANGVWRLSAPLPAFYQALDRRLCYQSMVLFFENYTG 60  
DB 1 MRYLLPSVLLGTAFTYVLANGVWRLSAPLPAFYQALDRRLCYQSMVLFFENYTG 60  
QY 61 VOILLYGDLPKKNENIYLANHSTVDMIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120

DB 61 VOILLYGDLPKKNENIYLANHSTVDMIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120  
QY 121 YFAOHGCIYVRSKAFNEKEKRNKLSQSYVDAGTMYLVIPEEGTRYNPEQTKVLSASQAF 180  
DB 121 YFAOHGCIYVRSKAFNEKEKRNKLSQSYVDAGTMYLVIPEEGTRYNPEQTKVLSASQAF 180  
QY 181 AAQGLAVLKHVLTPIKATHVAFDCKMKNYLDATYDVTVYEGKDDG 228  
DB 181 AAQGLAVLKHVLTPIKATHVAFDCKMKNYLDATYDVTVYEGKDDG 228

RESULT 9  
US-09-833-245-1474  
Sequence 1474, Application US/09833245  
GENERAL INFORMATION:  
APPLICANT: Human Genome Sciences, Inc.  
TITLE OF INVENTION: Albumin Fusion Proteins  
FILE REFERENCE: PF546PCT  
CURRENT APPLICATION NUMBER: US/09/833, 245  
CURRENT FILING DATE: 2001-04-12  
PRIOR APPLICATION NUMBER: 60/229, 358  
PRIOR FILING DATE: 2000-04-12  
PRIOR APPLICATION NUMBER: 60/256, 931  
PRIOR FILING DATE: 2000-12-21  
PRIOR APPLICATION NUMBER: 60/199, 384  
PRIOR FILING DATE: 2000-04-25  
NUMBER OF SEQ ID NOS: 2267  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 1474  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-833-245-1474

Query Match 100.0%; Score 1203; DB 22; Length 353;  
Best Local Similarity 100.0%; Pred. No. 1.4e-119;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTAFTYVLANGVWRLSAPLPAFYQALDRRLCYQSMVLFFENYTG 60  
DB 1 MRYLLPSVLLGTAFTYVLANGVWRLSAPLPAFYQALDRRLCYQSMVLFFENYTG 60  
QY 61 VOILLYGDLPKKNENIYLANHSTVDMIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120  
DB 61 VOILLYGDLPKKNENIYLANHSTVDMIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120  
QY 121 YFAOHGCIYVRSKAFNEKEKRNKLSQSYVDAGTMYLVIPEEGTRYNPEQTKVLSASQAF 180  
DB 121 YFAOHGCIYVRSKAFNEKEKRNKLSQSYVDAGTMYLVIPEEGTRYNPEQTKVLSASQAF 180  
QY 181 AAQGLAVLKHVLTPIKATHVAFDCKMKNYLDATYDVTVYEGKDDG 228  
DB 181 AAQGLAVLKHVLTPIKATHVAFDCKMKNYLDATYDVTVYEGKDDG 228

RESULT 10  
US-09-833-245-1475  
Sequence 1475, Application US/09833245  
GENERAL INFORMATION:  
APPLICANT: Human Genome Sciences, Inc.  
TITLE OF INVENTION: Albumin Fusion Proteins  
FILE REFERENCE: PF546PCT  
CURRENT APPLICATION NUMBER: US/09/833, 245  
CURRENT FILING DATE: 2001-04-12  
PRIOR APPLICATION NUMBER: 60/229, 358  
PRIOR FILING DATE: 2000-04-12  
PRIOR APPLICATION NUMBER: 60/256, 931  
PRIOR FILING DATE: 2000-12-21  
PRIOR APPLICATION NUMBER: 60/199, 384  
PRIOR FILING DATE: 2000-04-25  
NUMBER OF SEQ ID NOS: 2267

Query Match	100.0%; Score 1203; DB 22; Length 353;	Best Local Similarity 100.0%; Pred. No. 1,4e-119;	Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
NAME/KEY: CARBOHYD			
LOCATION: 57..59			
OTHER INFORMATION: Prosite match			
NAME/KEY: HELIX			
LOCATION: 76..96			
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method			
NAME/KEY: PHOSPHORYLATION			
LOCATION: 78			
OTHER INFORMATION: potential Tyrosine kinase site, Prosite match			
NAME/KEY: PHOSPHORYLATION			
LOCATION: 84			
OTHER INFORMATION: potential caseine kinase II site, Prosite match			
NAME/KEY: SITE			
LOCATION: 94..115			
OTHER INFORMATION: potential Leucine zipper site, Prosite match			
NAME/KEY: MYRISTATE			
LOCATION: 119..123			
OTHER INFORMATION: potential site, Prosite match			
NAME/KEY: PHOSPHORYLATION			
LOCATION: 133			
OTHER INFORMATION: potential protein kinase C, Prosite match			
NAME/KEY: PHOSPHORYLATION			
LOCATION: 147			
OTHER INFORMATION: potential caseine kinase II site, Prosite match			
NAME/KEY: PHOSPHORYLATION			
LOCATION: 194			
OTHER INFORMATION: potential protein kinase C, Prosite match			
NAME/KEY: PHOSPHORYLATION			
LOCATION: 215			
OTHER INFORMATION: potential Tyrosine kinase site, Prosite match			
NAME/KEY: SULFATATION			
LOCATION: 221			
OTHER INFORMATION: Prosite match			
NAME/KEY: PHOSPHORYLATION			
LOCATION: 233			
OTHER INFORMATION: potential cAMP and cGMP dependant protein kinase site,			
OTHER INFORMATION: match			
NAME/KEY: PHOSPHORYLATION			
LOCATION: 235			
OTHER INFORMATION: potential caseine kinase II site, Prosite match			
NAME/KEY: PHOSPHORYLATION			
LOCATION: 306			
OTHER INFORMATION: potential protein kinase C, Prosite match			
NAME/KEY: HELIX			
LOCATION: 310..330			
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method			
NAME/KEY: MYRISTATE			
LOCATION: 319..323			
OTHER INFORMATION: Prosite match			
NAME/KEY: MYRISTATE			
LOCATION: 323..327			
OTHER INFORMATION: Prosite match			
NAME/KEY: AMIDATION			
LOCATION: 329			
OTHER INFORMATION: Prosite match			
NAME/KEY: HELIX			
LOCATION: 333..353			
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method			
NAME/KEY: MYRISTATE			
LOCATION: 341..345			
OTHER INFORMATION: Prosite match			
NAME/KEY: PHOSPHORYLATION			
LOCATION: 350			
OTHER INFORMATION: potential protein kinase C, Prosite match			
US-09-853-526-4			

```
Qy 1 MRYLLPSVLLGTAFTYVLANGVWRLSAPLPARFYQALDRLCYCVOSWVLEFFENYTG 60
    |||
Db 1 MRYLLPSVLLGTAFTYVLANGVWRLSAPLPARFYQALDRLCYCVOSWVLEFFENYTG 60
Qy 61 VOILLYGDLPKNKENITLYLANHGSTVDMIVADIATRONALGHVRYVLEKGLKMLPLYGC 120
    |||
Db 61 VOILLYGDLPKNKENITLYLANHGSTVDMIVADIATRONALGHVRYVLEKGLKMLPLYGC 120
Qy 121 YFAOHCIGYVRSKAFENKRNKLSQSYVDAGTPLYVIEPEGTRYNPEQTKVLSASQAF 180
    |||
Db 121 YFAOHCIGYVRSKAFENKRNKLSQSYVDAGTPLYVIEPEGTRYNPEQTKVLSASQAF 180
Qy 181 AAQRCGLAVLKHVLPRIKATVAFEDCKMKNVLDATYDVTVYEGKDDG 228
    |||
Db 181 AAQRCGLAVLKHVLPRIKATVAFEDCKMKNVLDATYDVTVYEGKDDG 228

RESULT 12
US-09-901-484A-4
: Sequence 4, Application US/09901484A
: GENERAL INFORMATION:
: APPLICANT: Cohen, Daniel
: APPLICANT: Blumenfeld, Marta
: APPLICANT: Chumakov, Ilya
: APPLICANT: Bougueleret, Lydie
: TITLE OF INVENTION: Prostate Cancer Gene
: FILE REFERENCE: GEN-T11XC3D2
: CURRENT APPLICATION NUMBER: US/09/901,484A
: CURRENT FILING DATE: 2001-07-09
: PRIOR APPLICATION NUMBER: US 08/996,306
: PRIOR FILING DATE: 1997-12-22
: PRIOR APPLICATION NUMBER: US 60/099,658
: PRIOR FILING DATE: 1998-09-09
: PRIOR APPLICATION NUMBER: US 09/218,207
: PRIOR FILING DATE: 1998-12-22
: PRIOR APPLICATION NUMBER: US 09/338,907
: PRIOR FILING DATE: 1999-06-23
: PRIOR APPLICATION NUMBER: US 09/853,526
: NUMBER OF SEQ ID NOS: 578
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 4
: LENGTH: 353
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: HELIX
: LOCATION: (1)..(33)
: OTHER INFORMATION: Rao and Argos identification method, potential helix
: NAME/KEY: HELIX
: LOCATION: (4)..(20)
: OTHER INFORMATION: Klein, Kanehisa and Delisi identification method, potential helix
: NAME/KEY: HELIX
: LOCATION: (4)..(24)
: OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
: OTHER INFORMATION: potential helix
: NAME/KEY: LIPID
: LOCATION: (12)..(16)
: OTHER INFORMATION: MYRISTATE, Prosite match
: NAME/KEY: HELIX
: LOCATION: (50)..(70)
: OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
: OTHER INFORMATION: potential helix
: NAME/KEY: CARBOHYD
: LOCATION: (57)..(59)
: OTHER INFORMATION: Prosite match
: NAME/KEY: HELIX
: LOCATION: (76)..(96)
: OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
: OTHER INFORMATION: potential helix
: NAME/KEY: MOD_RES
: LOCATION: (78)..(78)
```

```
: OTHER INFORMATION: PHOSPHORYLATION, potential Tyrosine kinase site, Prosite match
: NAME/KEY: MOD_RES
: LOCATION: (84)..(84)
: OTHER INFORMATION: PHOSPHORYLATION, potential casein kinase II site, Prosite mat
: NAME/KEY: SITE
: LOCATION: (94)..(115)
: OTHER INFORMATION: Potential Leucine zipper site, Prosite match
: NAME/KEY: LIPID
: LOCATION: (119)..(123)
: OTHER INFORMATION: MYRISTATE, Prosite match
: NAME/KEY: MOD_RES
: LOCATION: (133)..(133)
: OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match
: NAME/KEY: MOD_RES
: LOCATION: (147)..(147)
: OTHER INFORMATION: PHOSPHORYLATION, potential casein kinase II, Prosite match
: NAME/KEY: MOD_RES
: LOCATION: (194)..(194)
: OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match
: NAME/KEY: MOD_RES
: LOCATION: (215)..(215)
: OTHER INFORMATION: PHOSPHORYLATION, potential Tyrosine kinase site, Prosite match
: NAME/KEY: MOD_RES
: LOCATION: (221)..(221)
: OTHER INFORMATION: SULFATATION, Prosite match
: NAME/KEY: MOD_RES
: LOCATION: (233)..(233)
: OTHER INFORMATION: PHOSPHORYLATION, potential cAMP and cGMP dependant protein kin
: NAME/KEY: MOD_RES
: LOCATION: (235)..(235)
: OTHER INFORMATION: PHOSPHORYLATION, potential casein kinase II site, Prosite mat
: NAME/KEY: MOD_RES
: LOCATION: (306)..(306)
: OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match
: NAME/KEY: HELIX
: LOCATION: (310)..(330)
: OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
: OTHER INFORMATION: potential helix
: NAME/KEY: LIPID
: LOCATION: (319)..(323)
: OTHER INFORMATION: MYRISTATE, Prosite match
: NAME/KEY: LIPID
: LOCATION: (323)..(327)
: OTHER INFORMATION: MYRISTATE, Prosite match
: NAME/KEY: MOD_RES
: LOCATION: (329)..(329)
: OTHER INFORMATION: AMIDATION, Prosite match
: NAME/KEY: HELIX
: LOCATION: (333)..(353)
: OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
: OTHER INFORMATION: potential helix
: NAME/KEY: LIPID
: LOCATION: (341)..(345)
: OTHER INFORMATION: MYRISTATE, Prosite match
: NAME/KEY: MOD_RES
: LOCATION: (350)..(350)
: OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match
US-09-901-484A-4
```

```
Query Match 100.0%; Score 1203; DB 23; Length 353;
Best Local Similarity 100.0%; Pred. No. 1,4e-119;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRYLLPSVLLGTAFTYVLANGVWRLSAPLPARFYQALDRLCYCVOSWVLEFFENYTG 60
    |||
Db 1 MRYLLPSVLLGTAFTYVLANGVWRLSAPLPARFYQALDRLCYCVOSWVLEFFENYTG 60
Qy 61 VOILLYGDLPKNKENITLYLANHGSTVDMIVADIATRONALGHVRYVLEKGLKMLPLYGC 120
    |||
Db 61 VOILLYGDLPKNKENITLYLANHGSTVDMIVADIATRONALGHVRYVLEKGLKMLPLYGC 120
```



[illegible]

```

RESULT 13
US-09-915-582-59
; Sequence 59, Application US/09915582
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 17 Human Secreted Proteins
; FILE REFERENCE: PS723p1
; CURRENT APPLICATION NUMBER: US/09/915,582
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: PCT/US01/01431
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/231,968
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-915-582-59

```

```

RESULT 14
US-09-915-582-79
Sequence 79, Application US/09915582
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 17 Human Secreted Proteins
FILE REFERENCE: PS723P1
CURRENT APPLICATION NUMBER: US/09/915,582
CURRENT FILING DATE: 2001-07-27
PRIORITY APPLICATION NUMBER: PCT/US01/01431
PRIORITY FILING DATE: 2001-01-17
PRIORITY APPLICATION NUMBER: 60/119,065
PRIORITY FILING DATE: 2000-01-31
PRIORITY APPLICATION NUMBER: 60/180,628
PRIORITY FILING DATE: 2000-02-04
PRIORITY APPLICATION NUMBER: 60/231,968
PRIORITY FILING DATE: 2000-09-12

```

```

:      NUMBER OF SEQ. ID NOS: 97
:      SOFTWARE: PatentIn Ver. 2.0
:      SEQ ID NO 79
:      LENGTH: 353
:      TYPE: prt
:      ORGANISM: Homo sapiens
:      OS-09-915-582-79

```

[illegible]

NAME/KEY: potential Transmembrane helix  
LOCATION: 1..33  
IDENTIFICATION METHOD: Rao and Argos method  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 4..20  
IDENTIFICATION METHOD: Klein, Kanehisa and Delisi method  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 4..24  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony  
IDENTIFICATION METHOD: and Wall method  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 12  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 50..70  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony and Wall method  
FEATURE:  
NAME/KEY: potential N-glycosylation site  
LOCATION: 57  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 76..96  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony and Wall method  
FEATURE:  
NAME/KEY: potential Tyrosine kinase phosphorylation site  
LOCATION: 78  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Caseln kinase II phosphorylation site  
LOCATION: 84  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Leucine zipper pattern  
LOCATION: 94..115  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 119  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 133  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Caseln kinase II phosphorylation site  
LOCATION: 147  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 194  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Tyrosine kinase phosphorylation site  
LOCATION: 215  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Tyrosine sulfatation site  
LOCATION: 221  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential cAMP- and cGMP-dependent protein kinase phosphorylation site  
LOCATION: 233  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Caseln kinase II phosphorylation site  
LOCATION: 235  
IDENTIFICATION METHOD: prosite match  
FEATURE:

NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 306  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 310..330  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony and Wall method  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 319  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 323  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Amidation site  
LOCATION: 329  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 333..353  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony and Wall method  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 341  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 350  
IDENTIFICATION METHOD: prosite match  
US-60-099-658-4

Query Match 100.0%; Score 1203; DB 26; Length 353;  
Best Local Similarity 100.0%; Pred. No. 1.4e-119;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MRYLLPSVVLGTPATYVLAMGWRLSAPLPARFYQALDRLCYQVSNVLEFFENYTG 60  
Db 1 MRYLLPSVVLGTPATYVLAMGWRLSAPLPARFYQALDRLCYQVSNVLEFFENYTG 60  
Qy 61 VQILLYGDLKRNKENTIIYLANHOSYVDIADILAIQNALGHVRYLKGKMLPLPGC 120  
Db 61 VQILLYGDLKRNKENTIIYLANHOSYVDIADILAIQNALGHVRYLKGKMLPLPGC 120  
Qy 121 YFAOHGSIYKRSKAFNEKEKRNKLOSVDAGTPMYLVIFPEGRYRNPEDTKVLSASQAF 180  
Db 121 YFAOHGSIYKRSKAFNEKEKRNKLOSVDAGTPMYLVIFPEGRYRNPEDTKVLSASQAF 180  
Qy 181 AAOGLAVLKHVLPRIKATHVAEDCMKNYLDIAYDTVYVEGKDDG 228  
Db 181 AAOGLAVLKHVLPRIKATHVAEDCMKNYLDIAYDTVYVEGKDDG 228

Search completed: August 28, 2002, 11:14:11  
Job time: 490 sec

---

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:15:42 ; Search time 94.08 Seconds  
(without alignments)  
656.370 Million cell updates/sec

Title: US-09-853-526-70  
Perfect score: 1203  
Sequence: 1 MRYLPSVILGTAFTYVLA.....NYLDATVDVTVVEGKDDGG 228

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 747981 seqs, 242050750 residues

Total number of hits satisfying chosen parameters: 747981

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1203	100.0	364	6	US-10-184-648-63
2	1200	99.8	353	5	US-09-629-469A-13028
3	759	63.1	269	6	US-10-074-045-47
4	291	24.2	375	5	US-09-935-625-11061
5	291	24.2	375	5	US-09-935-625-14197
6	291	24.2	375	5	US-09-935-625-14201
7	291	24.2	375	5	US-09-935-625-14254
8	291	24.2	375	5	US-09-935-625-23009
9	291	24.2	375	5	US-09-935-625-23025
10	288	23.9	351	5	US-09-935-625-11062
11	288	23.9	351	5	US-09-935-625-14198
12	288	23.9	351	5	US-09-935-625-14202
13	288	23.9	351	5	US-09-935-625-14255
14	288	23.9	351	5	US-09-935-625-23010
15	288	23.9	351	5	US-09-935-625-23026
16	288	23.9	375	5	US-09-935-625-11471
17	285	23.7	351	5	US-09-935-625-11472
18	282	23.4	400	7	US-60-391-781-1157
19	275	22.9	343	5	US-09-935-625-11063
20	275	22.9	343	5	US-09-935-625-11473
21	275	22.9	343	5	US-09-935-625-14199
22	275	22.9	343	5	US-09-935-625-14203
23	275	22.9	343	5	US-09-935-625-14256
24	275	22.9	343	5	US-09-935-625-23011
25	275	22.9	343	5	US-09-935-625-23027
26	233	19.4	414	6	US-10-121-062-314

27	233	19.4	414	6	US-10-063-502-102	Sequence 102, App
28	233	19.4	414	6	US-10-063-510-102	Sequence 102, App
29	233	19.4	414	6	US-10-063-512-102	Sequence 102, App
30	233	19.4	414	6	US-10-063-513-102	Sequence 102, App
31	233	19.4	414	6	US-10-063-514-102	Sequence 102, App
32	233	19.4	414	6	US-10-063-515-102	Sequence 102, App
33	233	19.4	414	6	US-10-063-516-102	Sequence 102, App
34	233	19.4	414	6	US-10-063-517-102	Sequence 102, App
35	233	19.4	414	6	US-10-063-518-102	Sequence 102, App
36	233	19.4	414	6	US-10-063-519-102	Sequence 102, App
37	233	19.4	414	6	US-10-063-520-102	Sequence 102, App
38	233	19.4	414	6	US-10-063-521-102	Sequence 102, App
39	233	19.4	414	6	US-10-063-522-102	Sequence 102, App
40	233	19.4	414	6	US-10-063-524-102	Sequence 102, App
41	233	19.4	414	6	US-10-063-525-102	Sequence 102, App
42	233	19.4	414	6	US-10-063-526-102	Sequence 102, App
43	233	19.4	414	6	US-10-063-527-102	Sequence 102, App
44	233	19.4	414	6	US-10-063-528-102	Sequence 102, App
45	233	19.4	414	6	US-10-063-529-102	Sequence 102, App

## ALIGNMENTS

RESULT 1  
US-10-184-648-63  
Sequence 63, Application US/10184648  
GENERAL INFORMATION:  
APPLICANT: Meyers, Rachel E.  
APPLICANT: Williamson, Mark  
APPLICANT: Tsai, Feng-Ying  
APPLICANT: Hunter, John J.  
APPLICANT: Macbeth, Kyle J.  
APPLICANT: Rudolph-Owen, Laura A.  
APPLICANT: Leiby, Kevin R.  
APPLICANT: Kapeller-Libermann, Rosana  
APPLICANT: Olandt, Peter J.  
TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE FAMILY MEMBERS AND USES THEREOF  
FILE REFERENCE: 10448-192001  
CURRENT APPLICATION NUMBER: US/10/184, 648  
CURRENT FILING DATE: 2002-06-27  
PRIOR APPLICATION NUMBER: US 09/815, 028  
PRIOR FILING DATE: 2001-03-22  
PRIOR APPLICATION NUMBER: PCT/US01/09358  
PRIOR FILING DATE: 2001-03-22  
PRIOR APPLICATION NUMBER: US 60/191, 964  
PRIOR FILING DATE: 2000-03-24  
PRIOR APPLICATION NUMBER: US 09/801, 220  
PRIOR FILING DATE: 2001-03-07  
PRIOR APPLICATION NUMBER: PCT/US01/07269  
PRIOR FILING DATE: 2001-03-07  
PRIOR APPLICATION NUMBER: US 60/187, 456  
PRIOR FILING DATE: 2000-03-07  
PRIOR APPLICATION NUMBER: US 09/816, 714  
PRIOR FILING DATE: 2001-03-23  
PRIOR APPLICATION NUMBER: PCT/US01/09468  
PRIOR FILING DATE: 2001-03-23  
PRIOR APPLICATION NUMBER: US 60/191, 865  
PRIOR FILING DATE: 2000-03-24  
PRIOR APPLICATION NUMBER: US 09/844, 948  
PRIOR FILING DATE: 2001-04-27  
PRIOR APPLICATION NUMBER: PCT/US01/13805  
PRIOR FILING DATE: 2001-04-27  
PRIOR APPLICATION NUMBER: US 60/200, 604  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: US 09/861, 164  
PRIOR FILING DATE: 2001-05-18  
PRIOR APPLICATION NUMBER: PCT/US01/16292  
PRIOR FILING DATE: 2001-05-18  
PRIOR APPLICATION NUMBER: US 60/205, 408  
PRIOR FILING DATE: 2000-05-19  
PRIOR APPLICATION NUMBER: US 09/883, 060  
PRIOR FILING DATE: 2001-06-15

```

; PRIOR APPLICATION NUMBER: PCT/US01/19138
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/2212,079
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/962,678
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: PCT/US01/29963
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/235,044
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 09/973,457
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: US 60/238,849
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 10/072,285
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: PCT/US02/03736
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/267,494
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: US 09/817,910
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: PCT/US01/09633
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/192,092
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: US 09/842,528
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: PCT/US01/40607
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 60/199,500
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/882,836
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19543
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/211,730
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/882,872
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19153
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,077
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-184-648-63

```

```

Query Match      100.0%; Score 1203; DB 6; Length 364;
Best Local Similarity 100.0%; Pred. No. 4,3e-116;
Matches 228; Conservatve 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MRYLLPSVLLGTAFTYVLANGVWRLLSAFLPARFYQALDRLKCVYQSMVLPFFENYTG 60
    |||||||
DB 12 MRYLLPSVLLGTAFTYVLANGVWRLLSAFLPARFYQALDRLKCVYQSMVLPFFENYTG 71
    |||||||
QY 61 VOILLYGDLPKKNENIYLANHSTVDIYADILAIRQNALGHVRYVLEKGLKMLPLYGC 120
    |||||||
DB 72 VOILLYGDLPKKNENIYLANHSTVDIYADILAIRQNALGHVRYVLEKGLKMLPLYGC 131
    |||||||
QY 121 YFAOHGIGIYVRSKAKFNEKEKRNKLOSYVDAGTSMYLVIPFEGTRYNPEQTKVLSASQAF 180
    |||||||
DB 132 YFAOHGIGIYVRSKAKFNEKEKRNKLOSYVDAGTSMYLVIPFEGTRYNPEQTKVLSASQAF 191
    |||||||
QY 181 AAQGLAVLKHVLTPIRIKATHVAFDCKKNYIDALYDTVYVEGKDDG 228
    |||||||
DB 192 AAQGLAVLKHVLTPIRIKATHVAFDCKKNYIDALYDTVYVEGKDDG 239
    |||||||

```

```

RESULT 2
US-09-629-469A-13028
; Sequence 13028, Application US/09629469A
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: SAITO, KAORU
; APPLICANT: YAMAMOTO, JUNICHI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: SUGITAMA, TOMOYASU
; APPLICANT: WAKAMATSU, AI
; APPLICANT: MAGAI, KEIICHI
; APPLICANT: OTSUKI, TETSUJI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE
; FILE REFERENCE: 084335/0123
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US/09/629,469A
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 1999-248036
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183,322
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 19025
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13028
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-629-469A-13028

```

```

Query Match      99.8%; Score 1200; DB 5; Length 353;
Best Local Similarity 99.6%; Pred. No. 8,4e-116;
Matches 227; Conservatve 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MRYLLPSVLLGTAFTYVLANGVWRLLSAFLPARFYQALDRLKCVYQSMVLPFFENYTG 60
    |||||||
DB 1 MRYLLPSVLLGTAFTYVLANGVWRLLSAFLPARFYQALDRLKCVYQSMVLPFFENYTG 60
    |||||||
QY 61 VOILLYGDLPKKNENIYLANHSTVDIYADILAIRQNALGHVRYVLEKGLKMLPLYGC 120
    |||||||
DB 61 VOILLYGDLPKKNENIYLANHSTVDIYADILAIRQNALGHVRYVLEKGLKMLPLYGC 120
    |||||||
QY 121 YFAOHGIGIYVRSKAKFNEKEKRNKLOSYVDAGTSMYLVIPFEGTRYNPEQTKVLSASQAF 180
    |||||||
DB 121 YFAOHGIGIYVRSKAKFNEKEKRNKLOSYVDAGTSMYLVIPFEGTRYNPEQTKVLSASQAF 180
    |||||||
QY 181 AAQGLAVLKHVLTPIRIKATHVAFDCKKNYIDALYDTVYVEGKDDG 228
    |||||||
DB 181 AAQGLAVLKHVLTPIRIKATHVAFDCKKNYIDALYDTVYVEGKDDG 228
    |||||||

```

RESULT 3  
US-10-074-045-47  
; Sequence 47, Application US/10074045  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT221C1  
; CURRENT APPLICATION NUMBER: US/10/074,045  
; CURRENT FILING DATE: 2002-02-14

```
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 47
; LENGTH: 269
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-074-045-87
```

```
Query Match          63.1%; Score 759; DB 6; Length 269;
Best Local Similarity 100.0%; Pred. No. 2.7e-70;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 85 TWDVIADILAIROMALGHVRYLKEGLKMLPIYGCYPAOHGSIYVRSKAFNEKEMRNK 144
DB 1 TYDVIADILAIROMALGHVRYLKEGLKMLPIYGCYPAOHGSIYVRSKAFNEKEMRNK 60
QY 145 LOSYVDAGTPMYLVIFPEGSTRYNPEQTKVLSASQAFPAQGLAVLKHLVLPRIKATHVAF 204
DB 61 LOSYVDAGTPMYLVIFPEGSTRYNPEQTKVLSASQAFPAQGLAVLKHLVLPRIKATHVAF 120
QY 205 DCMKNTLDIAYTVYVYEGKDDGG 228
DB 121 DCMKNTLDIAYTVYVYEGKDDGG 144
```

## RESULT 4

```
US-09-935-625-11061
; Sequence 11061, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 11061
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..375
; OTHER INFORMATION: Ceres Seq. ID no. 1394370
US-09-935-625-11061
```

```
Query Match          24.2%; Score 291; DB 5; Length 375;
Best Local Similarity 31.3%; Pred. No. 1.2e-21;
Matches 71; Conservative 50; Mismatches 94; Indels 12; Gaps 3;
```

```
QY 1 MRYLPSVVLGTAPTYVILAMG-----VWRLSAPLPARFYQALDRLXCVYOSMVLFFP 55
DB 18 LRGIICLMVLVSTAFMMLIFMGFLSAVLRIFS---IRYSRCVSPFFGSMWLMPLFLF 73
QY 56 ENYTGVOILLYGDLPKNKNIITYLANHOSTVDMIVADILAIROMALGHVRYLKEGLKWL 115
DB 74 EKINKTKVIFSGDKVPCEDRVLLIANHRTVEVDMYFWDLALRRGOIGNIKYVLSKSLMKL 133
QY 116 PLYGCYPAOHGSIYVRSKAFNEKEMRNKLOSVDAGTPMYLVIFPEGSTRYNPEQTKVLS 175
DB 134 PLFGMAFHLEFIPVERRMEVDANLRQIVSSFKDRPDALMLALFPEGIDTYTAKCO--- 190
QY 176 ASQAFPAQGLAVLKHLVLPRIKATHVAFDCMKNTLDIAYTVYVY 222
DB 191 RSKKFAENGLPILNNVLLPRTKGFVSCIQELSCSLDAVYDVITGYK 237
```

## RESULT 5

```
US-09-935-625-14197
; Sequence 14197, Application US/09935625
```

```
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 14197
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..375
; OTHER INFORMATION: Ceres Seq. ID no. 3441314
US-09-935-625-14197
```

```
Query Match          24.2%; Score 291; DB 5; Length 375;
Best Local Similarity 31.3%; Pred. No. 1.2e-21;
Matches 71; Conservative 50; Mismatches 94; Indels 12; Gaps 3;
```

```
QY 1 MRYLPSVVLGTAPTYVILAMG-----VWRLSAPLPARFYQALDRLXCVYOSMVLFFP 55
DB 18 LRGIICLMVLVSTAFMMLIFMGFLSAVLRIFS---IRYSRCVSPFFGSMWLMPLFLF 73
QY 56 ENYTGVOILLYGDLPKNKNIITYLANHOSTVDMIVADILAIROMALGHVRYLKEGLKWL 115
DB 74 EKINKTKVIFSGDKVPCEDRVLLIANHRTVEVDMYFWDLALRRGOIGNIKYVLSKSLMKL 133
QY 116 PLYGCYPAOHGSIYVRSKAFNEKEMRNKLOSVDAGTPMYLVIFPEGSTRYNPEQTKVLS 175
DB 134 PLFGMAFHLEFIPVERRMEVDANLRQIVSSFKDRPDALMLALFPEGIDTYTAKCO--- 190
QY 176 ASQAFPAQGLAVLKHLVLPRIKATHVAFDCMKNTLDIAYTVYVY 222
DB 191 RSKKFAENGLPILNNVLLPRTKGFVSCIQELSCSLDAVYDVITGYK 237
```

## RESULT 6

```
US-09-935-625-14201
; Sequence 14201, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 14201
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..375
; OTHER INFORMATION: Ceres Seq. ID no. 3441318
US-09-935-625-14201
```

```
Query Match          24.2%; Score 291; DB 5; Length 375;
Best Local Similarity 31.3%; Pred. No. 1.2e-21;
Matches 71; Conservative 50; Mismatches 94; Indels 12; Gaps 3;
```

```
QY 1 MRYLPSVVLGTAPTYVILAMG-----VWRLSAPLPARFYQALDRLXCVYOSMVLFFP 55
DB 18 LRGIICLMVLVSTAFMMLIFMGFLSAVLRIFS---IRYSRCVSPFFGSMWLMPLFLF 73
QY 56 ENYTGVOILLYGDLPKNKNIITYLANHOSTVDMIVADILAIROMALGHVRYLKEGLKWL 115
DB 74 EKINKTKVIFSGDKVPCEDRVLLIANHRTVEVDMYFWDLALRRGOIGNIKYVLSKSLMKL 133
```

QY 116 PLVGCYFAOHGIIYKRSANFENKEMRNKLSQSYVDAGTPLYLTFPEGSTRYNPEOTKVL 175  
Db 134 PLFGMAHLEFFIYVERKWEDEANLQIYSSFKDPRDALMLALFPEGTDTTEAKCQ--- 190  
QY 176 ASQAFAAORGLAVLKHVLTPIKATHVAFDCMKNYLDAIYDVYVE 222  
Db 191 RSKKFAENGPIILNNVLLPRTKGFVSCLOELSCSLDAVYDVDTIGYK 237

## RESULT 7

US-09-935-625-14254  
; Sequence 14254, Application US/09935625  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE  
; FILE REFERENCE: 2750-1481P  
; CURRENT APPLICATION NUMBER: US/09/935,625  
; NUMBER OF SEQ ID NOS: 2001-08-24  
; SEQ ID NO 14254  
; LENGTH: 375  
; TYPE: PRP  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: peptide  
; LOCATION: 1..375  
; OTHER INFORMATION: Ceres Seq. ID no. 3443248  
US-09-935-625-14254

Query Match 24.2%; Score 291; DB 5; Length 375;  
Best Local Similarity 31.3%; Pred. No. 1.2e-21;  
Matches 71; Conservative 50; Mismatches 94; Indels 12; Gaps 3;

QY 1 MRYLLPSVLLGTAPTYVLWG-----VWRLLSAFLPAREFYQALDRLCYVQSMVLEFF 55  
Db 18 LRGIICLMVIVSTAFMMLIFWGFSAVAVLRIFS---IRYSRKCVSFFGSMALMPFLF 73  
QY 56 ENTGVOIILLYGDLPRKKNENIYLANHOSYVDWIVADILAIROMALGHVRYLKEGLKWL 115  
Db 74 EKINKTVIISGDKVPCEDRVLLIANHRTEDWMEYDMLARKQIGINIKIVLSSLMKL 133  
QY 116 PLVGCYFAOHGIIYKRSANFENKEMRNKLSQSYVDAGTPLYLTFPEGSTRYNPEOTKVL 175  
Db 134 PLFGMAHLEFFIYVERKWEDEANLQIYSSFKDPRDALMLALFPEGTDTTEAKCQ--- 190  
QY 176 ASQAFAAORGLAVLKHVLTPIKATHVAFDCMKNYLDAIYDVYVE 222  
Db 191 RSKKFAENGPIILNNVLLPRTKGFVSCLOELSCSLDAVYDVDTIGYK 237

## RESULT 8

US-09-935-625-23009  
; Sequence 23009, Application US/09935625  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE  
; FILE REFERENCE: 2750-1481P  
; CURRENT APPLICATION NUMBER: US/09/935,625  
; NUMBER OF SEQ ID NOS: 33136  
; SEQ ID NO 23009  
; LENGTH: 375  
; TYPE: PRP  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: peptide  
; LOCATION: 1..375  
; OTHER INFORMATION: Ceres Seq. ID no. 3441318  
US-09-935-625-23009

Query Match 24.2%; Score 291; DB 5; Length 375;  
Best Local Similarity 31.3%; Pred. No. 1.2e-21;  
Matches 71; Conservative 50; Mismatches 94; Indels 12; Gaps 3;

QY 1 MRYLLPSVLLGTAPTYVLWG-----VWRLLSAFLPAREFYQALDRLCYVQSMVLEFF 55  
Db 18 LRGIICLMVIVSTAFMMLIFWGFSAVAVLRIFS---IRYSRKCVSFFGSMALMPFLF 73  
QY 56 ENTGVOIILLYGDLPRKKNENIYLANHOSYVDWIVADILAIROMALGHVRYLKEGLKWL 115  
Db 74 EKINKTVIISGDKVPCEDRVLLIANHRTEDWMEYDMLARKQIGINIKIVLSSLMKL 133  
QY 116 PLVGCYFAOHGIIYKRSANFENKEMRNKLSQSYVDAGTPLYLTFPEGSTRYNPEOTKVL 175  
Db 134 PLFGMAHLEFFIYVERKWEDEANLQIYSSFKDPRDALMLALFPEGTDTTEAKCQ--- 190  
QY 176 ASQAFAAORGLAVLKHVLTPIKATHVAFDCMKNYLDAIYDVYVE 222  
Db 191 RSKKFAENGPIILNNVLLPRTKGFVSCLOELSCSLDAVYDVDTIGYK 237

## RESULT 9

US-09-935-625-23025  
; Sequence 23025, Application US/09935625  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA  
; FILE REFERENCE: 2750-1481P  
; CURRENT APPLICATION NUMBER: US/09/935,625  
; NUMBER OF SEQ ID NOS: 33136  
; SEQ ID NO 23025  
; LENGTH: 375  
; TYPE: PRP  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: peptide  
; LOCATION: 1..375  
; OTHER INFORMATION: Ceres Seq. ID no. 3443248  
US-09-935-625-23025

Query Match 24.2%; Score 291; DB 5; Length 375;  
Best Local Similarity 31.3%; Pred. No. 1.2e-21;  
Matches 71; Conservative 50; Mismatches 94; Indels 12; Gaps 3;

QY 1 MRYLLPSVLLGTAPTYVLWG-----VWRLLSAFLPAREFYQALDRLCYVQSMVLEFF 55  
Db 18 LRGIICLMVIVSTAFMMLIFWGFSAVAVLRIFS---IRYSRKCVSFFGSMALMPFLF 73  
QY 56 ENTGVOIILLYGDLPRKKNENIYLANHOSYVDWIVADILAIROMALGHVRYLKEGLKWL 115  
Db 74 EKINKTVIISGDKVPCEDRVLLIANHRTEDWMEYDMLARKQIGINIKIVLSSLMKL 133  
QY 116 PLVGCYFAOHGIIYKRSANFENKEMRNKLSQSYVDAGTPLYLTFPEGSTRYNPEOTKVL 175  
Db 134 PLFGMAHLEFFIYVERKWEDEANLQIYSSFKDPRDALMLALFPEGTDTTEAKCQ--- 190  
QY 176 ASQAFAAORGLAVLKHVLTPIKATHVAFDCMKNYLDAIYDVYVE 222  
Db 191 RSKKFAENGPIILNNVLLPRTKGFVSCLOELSCSLDAVYDVDTIGYK 237

## RESULT 10

US-09-935-625-11062  
; Sequence 11062, Application US/09935625  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA  
; FILE REFERENCE: 2750-1481P  
; CURRENT APPLICATION NUMBER: US/09/935,625  
; NUMBER OF SEQ ID NOS: 33136  
; SEQ ID NO 11062  
; LENGTH: 375  
; TYPE: PRP  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: peptide  
; LOCATION: 1..375  
; OTHER INFORMATION: Ceres Seq. ID no. 3441318  
US-09-935-625-11062





```
Matches      70;    Conservative   47; Mismatches     91; Indels       12; Gaps        3;
Oy          8 VYLLGTAFTYYLVANG-----VRRLLSAPLPAFYQALDRLCYCQSNVLFFEFENITGYQ 62
           :|::||:::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db          1 MYLVSTAFEMMLIFMGFGLSAVLRLES---IRYSRKCVSFFFGSWLMALPFLFEKINKTK 56
Oy         63 ILLYGDLPKRKNENITYLVANHOSTVDIMIVADIILAIRONMAGHYRYULKEGKLWPLPGCYCF 122
           ::||::||::||::||::||||::||||::||::||::||::||::||::||::||::||::||
Db          57 YIFSCKDKPCCEBRLIINAHRTVEVDMWFMDLAKRGQIGNKIKYLKSLSMKLPLEGNMF 116
Oy         123 AQHGSIYVKRSARFKNEKEMRNKLQSYPVDAGTPMYLVIPSEGTGRYNPEQTGYVLSASQAEPAA 182
           ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db          117 HLFEEPIVERBRWEVDKANLRQLSVSFKDPRDALMWLALFPREGDYTEACQ--RSKKFAA 173
Oy         183 ORGLAVLKHVLTPTRIKAHVAFDCKMNVIDAIYDVTVAYIE 222
           :||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db          174 ENGLPILNNVLLPRTKGFSCLQELSCISDAVYDVTTIGK 213

RESULT_14
US-09-935-625-23010
: Sequence 23010, Application US/09935625
: GENERAL INFORMATION:
: APPLICANT: N. ALEXANDROV et al.
: TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE
: TITLE OF INVENTION: MODULATING VARIOUS RESPONSES
: FILE REFERENCE: 2750-1481P
: CURRENT APPLICATION NUMBER: US/09/935,625
: CURRENT FILING DATE: 2001-08-24
: NUMBER OF SEQ ID NOS: 33136
: SEQ ID NO 23010
: LENGTH: 351
: TYPE: PRT
: ORGANISM: Arabidopsis thaliana
: FEATURE:
: NAME/KEY: peptide
: LOCATION: 1..351
: OTHER INFORMATION: Ceres Seq. ID no. 3441319
US-09-935-625-23010
```

```

; SEQ ID NO 23026
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..351
; OTHER INFORMATION: Ceres Seq. ID no. 3443249
US-09-935-625-23026

Query Match          23.9%; Score 288; DB 5: Length 351;
Best Local Similarity 31.8%; Pred. No. 2.3e-21;
Matches 70; Conservative 47; Mismatches 91; Indels 12; Gaps 3

QY      8 VLLGTAFYYLVAMG-----VWRLSAFLPAFFYALDRLDKCYVQSMLEFFENYTGVO 62
       111: 11 : : 11 : 111 : : : : : : : : : : : : : : : : : : :
Db      1 MVLVSTAFEMMLIFMGFLSNVLRLS----IRSKCVSFFFGSLMALMPLEFKINKTK 56
       :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::

QY      63 ILLYGDLEPRKNENIITYLANHOSYVDMIVADILAFQNALGHNRYVYLEGSLKWLPLYGCF 122
       :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db      57 VIFSGDKVPCEDRVLILANHRTVEVDMMVFQMDLAKRGQGNIKYVLLKSSLMKLPLEGMAF 116
       :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::

QY      123 AOHGSIYYKRSKAFNEKEKRNKLGSYVDAGTMYLVIPEEGTRINYNQOTKYVLSASQAFPA 182
       111: 11 : : 11 : 111 : : : : : : : : : : : : : : : : : : :
Db      117 HLFEEIPIPERRWEVDENALROIYVSFFKDPDADALMLALEPFGTDYTEAKCO---RSRKFPA 173
       :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::

QY      183 QRGVLAVLKHVLTPRTIKATHVAFDCKKNLDAIYDVTYYVE 222
       :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
Db      174 ENGLPILNNVLLPRTKGVSVCLOELSCSDAIVAYDVTIGIK 213
       :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::

```

Search completed: August 28, 2002, 11:15:42  
Job time: 556 sec

[illegible]

---

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:08:58 ; Search time 39.13 Seconds

(without alignments)  
559.887 Million cell updates/sec

Title: US-09-853-526-70

Perfect score: 1203  
Sequence: 1 MRVLLPSTVLLGTAPTYVLA.....NYLDATDYVTVEGKDDGG 228

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	338	28.1	918	2 T34057	hypothetical prote
2	312	25.9	393	2 B96780	hypothetical prote
3	269	22.4	374	2 S52645	probable 1-acyl-gl
4	238	19.8	377	2 S60478	probable 1-acyl-sn
5	213.5	17.7	376	2 D96550	hypothetical prote
6	200	16.6	344	2 T31913	hypothetical prote
7	186.5	15.5	396	2 S54641	probable membrane
8	171	14.2	311	2 T07936	probable glycerol-
9	168.5	14.0	310	2 T06755	probable glycerol-
10	163.5	13.6	523	2 T25998	hypothetical prote
11	161.5	13.4	350	2 T40466	probable acetyltra
12	161	13.4	397	2 S45900	probable membrane
13	141	11.7	295	2 B83587	probable polynucle
14	139.5	11.6	285	2 T15252	hypothetical prote
15	134	11.1	439	2 T22689	hypothetical prote
16	124	10.3	302	2 A10950	probable acyltrans
17	120.5	10.0	303	2 A48600	probable sn2-acylg
18	120.5	10.0	310	2 S40808	polynucleotide ade
19	120.5	10.0	310	2 A91227	probable endonucle
20	120.5	10.0	310	2 G86073	probable endonucle
21	118	9.8	281	2 S60477	1-acylglycerol-3-p
22	117.5	9.8	391	2 T15366	hypothetical prote
23	114	9.5	266	2 S73863	probable 1-acylgly
24	113	9.4	247	2 G72223	hypothetical prote
25	112	9.3	304	2 G83101	probable acyltrans
26	112	9.3	304	2 B83541	probable polynucle
27	111	9.2	377	1 A44216	major envelope ant
28	109	9.1	240	2 D64089	1-acylglycerol-3-p
29	108	9.0	261	2 C82067	probable 1-acylgly

30	104	8.6	240	2 D64688	probable 1-acylgly
31	104	8.6	267	2 AC3499	1-acyl-sn-glycerol
32	103	8.6	243	2 AC0083	1-acylglycerol-3-p
33	103	8.6	266	2 B98283	hypothetical prote
34	103	8.6	266	2 AG3000	1-acyl-sn-glycerol
35	102	8.5	212	2 G85357	hypothetical prote
36	101	8.4	323	2 C83940	sugar ABC transpor
37	99.5	8.3	237	2 B71827	probable 1-acylgly
38	99	8.2	225	2 B45582	probable 1-acylgly
39	99	8.2	243	2 B71706	probable 1-acylgly
40	99	8.2	245	2 AE0888	1-acyl-glycerol-3-
41	98	8.1	379	2 E96842	hypothetical prote
42	97.5	8.1	262	2 T22599	1-acylglycerol-3-p
43	97.5	8.1	488	1 OOBPHS	alkaline exonuclea
44	97	8.1	245	2 S20460	1-acylglycerol-3-p
45	97	8.1	245	2 F91116	1-acyl-sn-glycerol

## ALIGNMENTS

```

RESULT
1
T34057
hypothetical protein F28B3.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T34057
R:Geisel, C.; Kramer, J.; Smith, A.
submitted to the EMBL Data Library, May 1997
A:Description: The sequence of C. elegans cosmid F28B3.
A:Reference number: Z21469
A:Accession: T34057
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-918 <GE1>
A:Cross-references: EMBL:AF003136; PIDN:AB93636.1; GSPDB:GN00019; CESP:F28B3.5
A:Experimental source: strain Bristol N2; clone F28B3
C:Genetics:
A:Gene: CESP:F28B3.5
A:Map position: 1
A:Introns: 85/3; 129/3; 235/3; 418/2; 482/3; 532/3; 736/3; 829/3; 850/2

Query Match      28 1%; Score 338; DB 2; Length 918;
Best Local Similarity 35.5%; Pred. No. 8.6e-23;
Matches 83; Conservative 48; Mismatches 85; Indels 18; Gaps 6;

QY      1 MRVLLP-----SVVLGTAPTYVLAAGVWRLSAFLPARFYQALDDRLCYQSMVLEFFE 56
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      544 LRPIPCSLSLSLSMVPPASCAIYIGV-----SMIVPRHVAQQLDMMLYKSMCLFVFE 599

QY      57 NYTVQVILLYGLDLPK--NK-----ENIIYLANHOSYVDWIVADILAIRONALGH---VRYV 107
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      600 NLGVEIYLHGTNEEVYVKNKPEKNAWMISNOSNVDMIIIPVMAARHGQGEQAFRYM 659

QY      108 LKSGIKMLPLYGCFYQHGSIYVRSKAFNEKMRNKLQSYVAGPFRMYVIREGRYV 167
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      660 VKNSIHLPWFPGWIFPHGIIYVRREGFTGAPVLRQLKMLNSDPPYLLIRPEGRFNS 719

QY      168 PEOTKVLASQAFAAQCGLAVLAKVHLPRIKATHVAFDCMKNYLDATYDVTYVY 221
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      720 AKKHILLESSNRFLEKSGRQPMQNVLCPRSGGQLALDNDNST-LDATYDVTYVY 772

RESULT
2
B96780
hypothetical protein F9E10.13 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: B96780
R:Theologian, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
Chn, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Hultzar, L.

```



```

Query Match          17.7%; Score 213.5; DB 2; Length 376;
Best Local Similarity 29.7%; Pred. No. 7.7e-12;
Matches 52; Conservative 46; Mismatches 66; Indels 11; Gaps 4;

OY 51 VLEFFENVTGVOILLYGDLPK---NKENIITLANHSTVDIMVADILAIRONALGHVY 106
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 57 LITLFDWMACIKINLYVDETLELIGKEHALVLSNHRSDIDLWIGVMAQRVCGSSIA 116

OY 107 VKEELKMLPLTG--CYFQHGIGYVKKRSKANEKEMRNKLOSYDAGCPMLVIFPECT 164
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 117 IMKKRAKTLPIITIGMSMWSDY--IFLEKSWAKDETLIKGFRLDEFPMTFMLALFEVET 174

OY 165 RYNEPOTKVLASQAFAAQRGLAVLKHVLTTPRIKATHVAFDCMKNYLDAYDTV 219
   ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 175 RTQGRK---LEAAQRYASIRSLPSPRNVLIPTKGFVSIVSELRSVPAIVDTCTL 226

RESULT 6
31913
hypothetical protein T05H4.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T31913
R:Blanchard, M.
Submitted to the EMBL Data Library, July 1997
A:Description: The sequence of C. elegans cosmid T05H4.
A:Reference number: 221097
A:Accession: T31913
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-344 <BLA>
A:Cross-references: EMBL:AF016452; PIDN:AA66008.1; GSPDB:GN00023; CESP:T05H4.1
A:Experimental source: strain Bristol N2; clone T05H4
A:Genetics:
A:Gene: CESP:T05H4.1
A:Map position: 5
A:Introns: 55/3; 212/1; 243/3; 298/3

Query Match          16.6%; Score 200; DB 2; Length 344;
Best Local Similarity 28.7%; Pred. No. 1.2e-10;
Matches 66; Conservative 35; Mismatches 93; Indels 36; Gaps 10;

OY 10 ILGTG----PTVLANGVRLISAFLEPARFYQALDDRLCYQSMWLFEEYNTGVQIL 65
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 18 LLGTFLFLFLPLPLAW-----FAP-KLWRTCADRLVGFWDITPCSLIEVYFGVNFY 68

OY 66 YGDLPRKNENIITLANHSTVDIMVADILAIRONALGHV-----RYLKEGILKMLPL 117
   ||| : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 69 TGDILRDEPALILMNHRTLDWLS-----WNALYKMDPMLLTTEKISTKAPLKKIP- 121

OY 118 YGCTYAGHGIGYVKKSAF-NEKEMRNKLOSYDAGCTPMY-LVIFPEGGRYNEPOTKVL 175
   : : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 122 -GAGNAMSGSYIFLDRNFENDKPVLERIVKYSGSEKKYQILFLAEGDKGERATRL-- 178

OY 176 ASQAFAAQRGLAVLKHVLTTPRIKATHVAFDCMK--NYLDAYDTVYVEG 223
   ||| : | : | : | : | : | : | : | : | : | : | : | : |
Db 179 -SDAFAADKNGLPREYEVLPHTTGKFKFLMELMKKENIYIVDTLTAISG 227

RESULT 7
55461
probable membrane protein YDR018c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein D3246; hypothetical protein PZP396; hypothetical
C:Species: Saccharomyces cerevisiae
C:Date: 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 20-Jun-2000
C:Accession: S54641; S63425; S67831; S72116
R:Pedman, K.; Brown, D.; Hamlyn, N.; Bowman, S.
Submitted to the EMBL Data Library, May 1995
A:Reference number: S54638
A:Accession: S54641
A:Molecule type: DNA
A:Residues: 1-396 <DED>

```

```

A:Cross-references: EMBL:Z49770; NID:g840867; PIDN:CAA89843.1; PID:g840871
A:Experimental source: strain AB972
R:Eide, L.G.; Sander, C.; Prydz, H.
Submitted to the EMBL Data Library, February 1996
A:Description: Sequencing and analysis of a 35.4 kb region on the left arm of chromos
A:Reference number: S63416
A:Accession: S63425
A:Molecule type: DNA
A:Residues: 1-396 <EID>
A:Cross-references: EMBL:X95966; NID:g1216215; PIDN:CAA65210.1; PID:g1216225
R:Prydz, H.; Eide, L.G.
Submitted to the Protein Sequence Database, July 1996
A:Reference number: S67822
A:Accession: S67831
A:Molecule type: DNA
A:Residues: 1-396 <PRY>
A:Cross-references: EMBL:Z74314; NID:g1431443; PIDN:CAA98838.1; PID:g1431444; MIPS:YD
A:Experimental source: strain S288C
R:Eide, L.G.; Sander, C.; Prydz, H.
Yeast 12. 1085-1090, 1996
A:Title: Sequencing and analysis of a 35.4 kb region on the left arm of chromosome IV
A:Reference number: S72107; MUID:97051598
A:Accession: S72116
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-396 <EIM>
A:Cross-references: EMBL:X95966; NID:g1216215; PIDN:CAA65210.1; PID:g1216225
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1996
C:Genetics:
A:Map position: 4R
A>Note: YDR018c
C:Superfamily: probable membrane protein YBR042c
C:Keywords: transmembrane protein
E:27-43/Domain: transmembrane #status predicted <TM1>
E:69-85/Domain: transmembrane #status predicted <TM2>
E:376-392/Domain: transmembrane #status predicted <TM3>

Query Match          15.5%; Score 186.5; DB 2; Length 396;
Best Local Similarity 31.0%; Pred. No. 2.5e-09;
Matches 53; Conservative 26; Mismatches 67; Indels 25; Gaps 4;

OY 73 KENIITLANHSTVDIMVADILAIRONALGHVRYVLEKGLKPLTYGCFYFAHGGIYVYR 132
   | : | : | : | | | | | | : | : | : | : | : | : | : |
Db 108 KDRATILANHQTADMIYIMLISFVSNLGGNYIILAKKLQIIFPLGFGMRNFKFTFLSR 167

OY 133 SAKFENEKERNKL-----QSYVDAGTPMY-LVIFPEGGRYNEPOTK 172
   : : | : | | | | | | | | | | | | | | | | | |
Db 168 MNQKDEKALTNLSVMDLNAKCGPLTNKSCYSKTNESIAVNLIMPEGTNLS---LK 224

OY 173 VLSAQAFAAQRGL--AVLKHVLTTPRIKATHVAFDCMKRYLDAYDTVYVY 221
   : : | : | : | : | : | : | : | : | : | : | : |
Db 225 TREKSEAFQRAHLDHVQRLHLPLPSKGLKRAVEKLAPSLDAIYDTVIGY 275

RESULT 8
T07936
probable glycerol-3-phosphate O-acetyltransferase (EC 2.3.1.15) - rape
N:Alternate names: 1-acyl-sn-glycerol-3-phosphate acyltransferase
C:Species: Brassica napus (rape)
C:Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 20-Jun-2000
C:Accession: T07936
R:Brough, C.L.
Submitted to the EMBL Data Library, June 1995
A:Reference number: Z16230
A:Accession: T07936
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-311 <BRO>
A:Cross-references: EMBL:Z49860; NID:g1149594; PIDN:CAA90019.1
A:Experimental source: cv. jet neuf; embryo
C:Superfamily: probable membrane protein YBR042c
C:Keywords: acyltransferase; coenzyme A

```

```

Query Match Similarity 14.2%; Score 171; DB 2; Length 311;
Best Local Similarity 32.4%; Pred. No. 4.8e-08;
Matches 44; Conservative 34; Mismatches 50; Indels 8; Gaps 4;

OY 86 VDMIVADILAI RQNALGHVRYLKEGLKMLPLYG--CYFAOHGIIYKRSKAFNEKEMRN 143
   : : : : : | | | : : : : : : : : : : : : : : : : : :
DB 17 ISGLVYNILIQ--RSGCIGSALAVWKSSKFLPIVIGSMFSEY--LFLERNMAKDESTLKS 73
   : : : : : | | | : : : : : : : : : : : : : : : : : :

OY 144 KIQSVDACTPMYLYITFPEGTRYNPEQRTKIVLSASQAFPAORGLAVLKHYLTPIRIATHTVA 203
   | | | | | : : : : : | | | | : : : : : | | | | | |
DB 74 GIQRINDPEPRPFWLLLFEGTRFTEAK--LKAQDEYAASSELLPVRNVLIPRTKGFVSA 130
   | | | | | : : : : : | | | | : : : : : | | | | | |

OY 204 FDCMKNYLDAIDYTV 219
   | : : : : | | | | |
DB 131 VSNMRSFVPALIDMTV 146
   | : : : : | | | | |

RESULT 9
T06755
Alternative glycerol-3-phosphate O-acyltransferase (EC 2.3.1.15) - Arabidopsis thaliana
P:Alternate names: protein F15B8.160
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 19-May-2000
C:Accession: T06755
R:Queller, F.; Benes, V.; Rechmann, S.; Borkova, D.; Ansoerge, W.; Salanoubat, M.; Mewes
  Submitted to the Protein Sequence Database, April 1999
A:Reference number: Z15794
A:Accession: T06755
A:Molecule type: DNA
A:Residues: 1-310 <QUE>
A:Cross-references: EMBL:AL049660; GSPDB:GN00061; ATSP:F15B8.160
A:Experimental source: cultivar Columbia; BAC clone F15B8
C:Genetics:
A:Gene: ATSP:F15B8.160
A:Map position: 3
A:Interons: 26/3; 46/3; 72/3; 125/3; 167/3; 198/3; 227/3
C:Superfamily: probable membrane protein YBR042c
C:Keywords: acyltransferase; coenzyme A

Query Match 14.0%; Score 168.5; DB 2; Length 310;
Best Local Similarity 32.8%; Pred. No. 8.2e-08;
Matches 41; Conservative 31; Mismatches 46; Indels 7; Gaps 3;

OY 97 RQNALGHVRYLKEGLKMLPLYG--CYFAOHGIIYKRSKAFNEKEMRNKIQSVYDAGTP 154
   | | | | : : : : | : : : : : : : : : : : : : : |
DB 27 RSGCIGSALAVWKSSKFLPIVIGSMFSEY--LFLERNMAKDESTLKSGLRLSDPEPR 84
   : : : : : | | | : : : : : : : : : : : : : : : : : :

OY 155 MYLVFPEGTRYNPEQRTKIVLSASQAFPAORGLAVLKHYLTPIRIATHTVAFDCMKNYLDAI 214
   : : : : : | | | : : : : : | | | | : : : : : | |
DB 85 FWLALFVEGTRFTEAK--LKAQDEYAASSELLPVRNVLIPRTKGFVSAVSNMRSFVPAL 141
   : : : : : | | | : : : : : | | | | : : : : : | |

OY 215 YDVTV 219
   | | | |
DB 142 YDMTV 146
   | | | |

RESULT 10
T25998
hypochemical protein ZK40.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T25998
R:Bradshaw, H.
  Submitted to the EMBL Data Library, July 1996
A:Description: The sequence of C. elegans cosmid ZK40.
A:Reference number: Z20121
A:Accession: T25998
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-523 <BRA>

```

```

A:Cross-references: EMBL:U64839; PIDN:AA804844.1; GSPDB:GN00023; CESP:ZK40.1
A:Experimental source: strain Bristol N2; clone ZK40
C:Genetics:
A:Gene: CESP:ZK40.1
A:Map position: 5
A:Introns: 7/1; 38/3; 93/3; 126/3; 160/1; 301/3; 422/2; 472/3

Query Match      13.6%; Score 153.5; DB 2; Length 523;
Best Local Similarity 25.4%; Pred. No. 4.6e-07;
Matches 57; Conservative 40; Mismatches 90; Indels 37; Gaps 5;

OY 6 PSVLLGTRAPTVVLAAGVWRLLSAFLPAPFYGALDRLYCVQSNVLFPEFFNYTGVOILL 65
DB 167 PLVLLL-----FLKPLRMQMDRLGLVIMPGALCNITFANIRI 208
OY 66 YDDLPRKKNENIYLANHSTVDVIMADILAIQNALGHV-----RYLKEKGLMPL 117
DB 209 KGDFLINHEPRLAIIMNHRIRLDWL-----FFWNALIKMDPMLCTTERISLKGMLKYPG 262
OY 118 YGCFYAGHGIGYVKKSAFNEKEMRNKLQSYVDAGTPMYLVIFPEGTRYNPEQTKVLSAS 177
DB 263 AGMAQAASAYIFLDRSFDTDKRLDNLINLYAYETEKYQOLLFPEGTDKCPKATE---RS 319
OY 178 QAFPAQRGIAVLKHYLTPIRIKATHVAFDCKM--NVLDAIYDVTY 219
DB 320 RIHSERKGLVHYOYVLPHPVTGFVHIYQAMRRANNIKRYIDVSI 363

RESULT 11
T40466
probable acetyltransferase protein - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T40466
R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Baker, S.; Mungall, K.
submitted to the EMBL Data Library, November 1998
A:Reference number: Z21931
A:Accession: T40466
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-350 <LNN>
A:Cross-references: EMBL:AL034382; PIDN:CAA22289.1; GSPDB:GN00067; SPDB:SPBC428.14
A:Experimental source: strain 97Zh-; cosmid c428
C:Genetics:
A:Gene: SPDB:SPBC428.14
A:Map position: 2

Query Match      13.4%; Score 161.5; DB 2; Length 350;
Best Local Similarity 27.7%; Pred. No. 4.2e-07;
Matches 41; Conservative 30; Mismatches 74; Indels 3; Gaps 2;

OY 74 ENIIYLANHQSVDVIMADILAIQNALGHVRYVLYKEGLKMWPLTGCFYAGHGIGYVRS 133
DB 94 ENIIYIANHQLYSDWMYVWMLSYTAKOHGHVITMLKSLIKMLPVLIGMGQMLFRFTFLSR 153
OY 134 AKFNEKEMRNKLQSYVDAGTPMYLVIFPEGTRYNPEQTKVLSASQAQFARGLAVLKHYL 193
DB 134 WDKDYEYHSRHKRFLRNVRDSYSLLPFGENTL-VEST--YQSRKYVADKIGVKKPKHLM 210
OY 194 TPRIKATHVAFDCKMKNVLDIAIYDVTY 221
DB 211 LPRVGRGLFYISISQLRDSMTYLYDTYTFY 238

RESULT 12
S45900
probable membrane protein YBR042c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein YBR0412
C:Species: Saccharomyces cerevisiae
C:Date: 26-Aug-1994 #sequence_revision 09-Sep-1994 #text_change 22-Oct-1999
C:Accession: S45900

```



R.Andre, B.; Cziepluch, C.; Hein, C.; Jauniaux, J.C.; Urrestarazu, A.; Vissers, S.  
submitted to the Protein Sequence Database, August 1994

A:Reference number: S45893

A:Accession: S45900

A:Molecule type: DNA

A:Residues: 1-397 <AND>

A:Cross-references: EMBL:235911; NID:g536265; PID:CAA84984.1; PID:g536266; GSPDB:GN0000

A:Experimental source: strain S288C

C:Genetics:

A:Gene: MIPS:YBR042C

A:Map position: 2R

C:Superfamily: probable membrane protein YBR042C

C:Keywords: transmembrane protein

F:12-37/Domain: transmembrane #status predicted <TM1>

F:55-77/Domain: transmembrane #status predicted <TM2>

F:134-150/Domain: transmembrane #status predicted <TM3>

F:372-390/Domain: transmembrane #status predicted <TM4>

#### Query Match

13.4%; Score 161; DB 2; Length 397;

Best Local Similarity 24.5%; Pred. No. 5.5e-07;

Matches 54; Conservative 35; Mismatches 89; Indels 42; Gaps 5;

QY 43 LYCVYQSWLFFPENYTGVOILLYGDPKN-----KENIYLANHQSVDWIVADILAI 96

DB 67 LHVVASAVRITTESSVPRKTEFDLKKRRLSHLKSNSVAICNHQIYTDWIFLWMLAV 126

QY 97 RQNALGHVRYVLEKQKMLPIYGCYFAQHGIVYKRSKAFNEKEMKNL----- 145

DB 127 TSNLGNVYLLKSLASIPILGFGKRNKVFPMRSKMAODKTTLSNLSLADGDSNRGAG 186

QY 146 -----QSYVDAGT-----PMYLVFPEGTRYPNQTKVLSASQAFPAQ 183

DB 187 SLAGKSPERITEGESSIMNPVIDPKQIHMPYLLIFPEGTINLSAD--TRQKSAK--YAAK 243

QY 184 KGLAVLKHLVTPRIKATVHAFDCMKNYLDALIDVTIVYRG 223

DB 244 IGKKPKFNVLPHSTGLRSLQKLKPSIESLYDITIGSG 283

#### RESULT 13

B83587

Probable polynucleotide adenyllyltransferase (EC 2.7.7.19) II PA0461 [similarity] - Pseud

C:Species: Pseudomonas aeruginosa

C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 02-Mar-2001

C:Accession: B83587

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lam,

.. Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

A:Reference number: A82950; MUID:20437337

A:Accession: B83587

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-295 <STO>

A:Cross-references: GB:AE004484; GB:AE004091; NID:g9946320; PIDN:AA603850.1; GSPDB:GN001

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA0461

C:Superfamily: Escherichia coli polynucleotide adenyllyltransferase II yIing

C:Keywords: nucleotidyltransferase

#### Query Match

11.7%; Score 141; DB 2; Length 295;

Best Local Similarity 26.4%; Pred. No. 2.6e-05;

Matches 64; Conservative 45; Mismatches 79; Indels 54; Gaps 13;

QY 4 LPSVVLGTAPRYVLANGVMLLSFLPARFQALDDRLCYQSWLFFPENYTGVOI 63

DB 15 LNTVTLTIG--PMMLIA-----LTKVLPLPGQ-----RLKDCSGKGVMIATWAEIDK 60

QY 64 LLY-----GDLPKNKENIYLANHQSVDWIVADILAIQNALGHVRYVLE 110

DB 61 WIPATCTPTVWDVRCASDLNRODTSYLV--VSNHQSVD--IPALIQVFNRRTPYKFFLKK 117

QY 111 GLKWLPLGYCYFAQHGIVYKRSKAF--FNKK--EMRKKLSQYVDAG-----TFMYLVIRP 161

DB 118 ELTWVFLGLAWMALEYPEKKRYSKAFLEKHPBLKGLDLETTAKACEKRRRIPIVTVNNYL 177

QY 162 EGRVYRPEQKTVLSASQAFNAQGLAVLKHLVTPRIKATVAF--DCMKNYLDAIVDVTY 219

DB 178 EGRTRFS--EQKKA-----RQSPYRHLKPF--KAGVATVLAALGEQDLALDVTY 224

QY 220 VY 221

DB 225 VY 226

#### RESULT 14

T15252

hypothetical protein K07B1.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999

C:Accession: T15252

R:Pauley, A.; Galtung, S.

submitted to the EMBL Data Library, May 1997

A:Description: The sequence of C. elegans cosmid K07B1.

A:Reference number: Z18317

A:Accession: T15252

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-285 <PAU>

A:Cross-references: EMBL:AF003384; NID:g2088817; PID:g2088819; PIDN:AA54237.1; GSPDB:

A:Experimental source: strain Bristol N2; clone K07B1

A:Gene: CESP:K07B1.5

A:Map position: 5

A:introns: 17/2; 72/2; 220/2

#### Query Match

11.6%; Score 139.5; DB 2; Length 285;

Best Local Similarity 24.2%; Pred. No. 3.4e-05;

Matches 44; Conservative 37; Mismatches 82; Indels 19; Gaps 8;

QY 34 RFYQALDDRLCYQSWLFFPENYT--GVQILLYGD---LPKNKENIYLANHQSVDWT 89

DB 74 RLTFWFBGKRYRLQSFIAFW--GYTAGYVYEGDVTYTYNDERVLMCNHQSTAD-- 129

QY 90 VADILAIKRON--ALGHVRYVLEKGLKWLPLCYCYFAQHGIVYKRSKAFNEKE--MRN 143

DB 130 VPTLMITVLQNKGVASRKTLMMDVMPFRWP--FGIIGNHGDYPIQOGKATRDRELRLKK 188

QY 144 KLOSIVDAGTPMVLVIFPEGTRYPNQTKVLSASQAFNAQGLAVLKHLVTPRIKATV 203

DB 189 HLHDVFWDRDRRWVILFPEGGFY----YKRVESSQSYGKRNKGPPLLTYTLPPKMGAVQAI 244

QY 204 FD 205

DB 245 LE 246

#### RESULT 15

T22689

hypothetical protein F55A11.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T22689

R:Kershaw, J.

submitted to the EMBL Data Library, May 1996

A:Reference number: Z19600

A:Accession: T22689

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-439 <WIL>

A:Cross-references: EMBL:Z72511; PIDN:CAA96659.1; GSPDB:GN00023; CESP:F55A11.5

A:Experimental source: clone F55A11  
C:Gene: CESP:F55A11.5  
A:Map position: 5  
A:Introns: 19/3; 52/2; 176/1; 222/3; 260/2; 327/1; 342/1

Query Match 11.1%; Score 134; DB 2; Length 439;  
Best Local Similarity 23.6%; Pred. No. 0.00019;  
Matches 45; Conservative 40; Mismatches 74; Indels 32; Gaps 4;  
QY 60 GVOILLYGDLPKKNKENTITLANQSTVDWIV--ADILAIRONALGHVRYVLEKGLKWLPL 117  
Db 27 GVRIRVSGDEIERGSPAMIVMNHRTLDWMYMCALYQINPWLITSNKISLKAQLKRLPG 86  
QY 118 YGCFYAOHGSIYVKRSKFNKEKEMRNKLOSVDAGTPMVLVIFPEGTRVYNPEOTKVLVSAS 177  
Db 87 AGFGMAAAGVFLERNAEVDKRSFDDAIDYFKNIDKKYQILLFPEGTDKSEWTT---LKS 143  
QY 178 QAFPAQOGLAVLKHVLTFR-----IKATHVAF-----DCMKNY 210  
Db 144 REFPAKKNGLRHLDYVLYPRTTGLHLNKMREPSSVIRFKNVHFYRFIKIQKIIILEEY 203  
QY 211 LDAIYDVTVYV 221  
Db 204 VEXIYDITIAV 214

Search completed: August 28, 2002, 11:08:59  
Job time: 253 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:17:23 ; Search time 21.99 Seconds

(without alignments)  
401.458 Million cell updates/sec

Title: US-09-853-526-70

Perfect score: 1203

Sequence: 1 MRYLPSVVLGTAPTYVLA.....NYLDAIVTVYEKGDDG 228

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1203	100.0	353	1 PLCE_HUMAN	Q9nqz2 homo sapien
2	1040	86.5	354	1 PLCE_MOUSE	Q9dl8r homo musculi
3	197.5	16.4	378	1 PLCD_RAT	Q92s1 ratu
4	196	16.3	378	1 PLCD_HUMAN	Q9nrt5 homo sapien
5	187.5	15.6	376	1 PLCC_HUMAN	Q9nrt7 homo sapien
6	186.5	15.5	396	1 YD18_YEAST	Q12185 saccharomyc
7	161	13.4	397	1 YBP2_YEAST	P38226 saccharomyc
8	125.5	10.4	370	1 Y205_HUMAN	Q92604 homo sapien
9	120.5	10.0	303	1 PLSC_YEAST	P33333 saccharomyc
10	120.5	10.0	310	1 YIHG_ECOLI	P32129 escherichia
11	119	9.9	281	1 PLSC_LIMPO	Q42870 limnanthes
12	117.5	9.8	391	1 Y773_CAEEL	Q11087 caenorhabdi
13	116	9.6	281	1 PLSC_LIMAL	Q42868 limnanthes
14	114	9.5	266	1 PLSC_MYCPN	P75479 mycoplasma
15	111	9.2	377	1 VENV_FOWPV	P36316 fowlpox vir
16	109	9.1	240	1 PLSC_HAELN	P44848 haemophilus
17	104	8.6	240	1 PLSC_COCNU	Q25903 helicobacte
18	101	8.4	308	1 PLSC_HELPI	Q42670 cocco
19	99.5	8.3	237	1 PLSC_HELPI	Q921n8 helicobacte
20	99	8.2	245	1 PLSC_SALTY	Q93471 salmonella
21	97.5	8.1	262	1 PLCL_CAEEL	Q93441 caenorhabdi
22	97.5	8.1	488	1 EXON_HSV6U	P24447 human herpe
23	97	8.1	245	1 PLSC_ECOLI	Q22267 escherichia
24	94.5	7.9	282	1 PLIC2_CAEEL	Q22267 caenorhabdi
25	94.5	7.9	488	1 EXON_HSV6Z	P52448 human herpe
26	93	7.7	250	1 PLSC_HSV6Z	Q59188 borrelia bu
27	90	7.5	484	1 Y142_NPVOF	O10370 oryza pseu
28	87.5	7.3	446	1 ML64_MOUSE	O61542 mus musculi
29	83.5	6.9	268	1 PLSC_MYCGE	Q24902 mycoplasma
30	82	6.8	503	1 CPN1_PAPHA	Q25527 papio hama
31	80	6.7	529	1 YOL1_CAEEL	P30638 caenorhabdi
32	79.5	6.6	2156	1 RRP1_PUDMH	P27176 pumila vir
33	79	6.6	281	1 Y7XQ_BACSU	P06569 bacillus su

34	78.5	6.5	204	1 NDCM_RECAM	O21271 reclinomona
35	78	6.5	445	1 ML64_HUMAN	Q14849 homo sapien
36	78	6.5	503	1 CPN1_HUMAN	P15538 homo sapien
37	77.5	6.4	371	1 Y1B0_CLOAB	O04354 clostridium
38	77.5	6.4	383	1 FDEE_ARATH	P46313 arabidopsis
39	77	6.4	999	1 DSG3_HUMAN	P32926 homo sapien
40	76.5	6.4	283	1 PLCA_HUMAN	Q99943 homo sapien
41	76.5	6.4	531	1 UD16_MOUSE	O64435 mus musculi
42	76	6.3	307	1 SYGA_XYLLA	O9PC25 xyella fas
43	76	6.3	384	1 FDEE_BRAUO	Q39287 brassica ju
44	76	6.3	892	1 RA16_SCHPO	P36617 schizosacch
45	75.5	6.3	326	1 305B_RAT	P31210 rattus norv

## ALIGNMENTS

RESULT	ID	PLCE_HUMAN	STANDARD	PRT	353 AA.
AC	Q9NIO2: Q9B0G4:				
DT	01-MAR-2002 (Rel. 41, Created)				
DT	01-MAR-2002 (Rel. 41, Last sequence update)				
DT	01-MAR-2002 (Rel. 41, Last annotation update)				
DE	1-acyl-sn-glycerol-3-phosphate acyltransferase epsilon (EC 2.3.1.51)				
DE	(1-AGP acyltransferase 5) (1-AGPAT 5) (Lysophosphatidic acid				
DE	acyltransferase-epsilon) (LPAAT-epsilon) (1-acylglycerol-3-phosphate				
DE	O-acyltransferase 5).				
GN	AGPAT5.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OX	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RP	SEQUENCE FROM N.A.				
RA	Leung D.W.;				
RT	"Cloning and expression of LPAAT-epsilon.";				
RL	Submitted (MAY-2001) to the EMBL/GenBank/DDJ databases.				
RN					
RP	SEQUENCE FROM N.A.				
RA	Cohen D., Chumakov I., Blumenfeld M., Bougueleret L.;				
RL	Patent number WO9332644, 01-JUL-1999.				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RA	TISSUE=amygdala;				
RX	MEDLINE=21154917; PubMed=11230166;				
RA	Wiemann S., Well B., Wellenreuther R., Gassenhuber J., Glassl S.,				
RA	Ansoerge W., Boecker M., Bloeker H., Bauersachs S., Blum H.,				
RA	Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,				
RA	Mewes H.-W., Oltmawelder B., Obermayer B., Tampe J., Heubner D.,				
RA	Wambolt R., Korn B., Klein M., Poustka A.;				
RT	"Towards a catalog of human genes and proteins: sequencing and				
RT	analysis of 500 novel complete protein coding human cDNAs.";				
RT	Genome Res. 11:422-435(2001).				
RP	SEQUENCE FROM N.A.				
RA	TISSUE=Placenta;				
RC					
RA	Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,				
RA	Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,				
RA	Wagatsuna M., Hosokiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,				
RA	Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,				
RA	Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,				
RA	Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y.,				
RA	Ninomiya K., Iwayanagi T.;				
RT	"NEO human cDNA sequencing project.";				
RT	Submitted (FEB-2000) to the EMBL/GenBank/DDJ databases.				
CC	-1- FUNCTION: CONVERTS LYSOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC				
CC	ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION (BY				
CC	SIMILARITY).				
CC	-1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate -				
CC	CoA + 1,2-diacyl-sn-glycerol 3-phosphate.				
CC	-1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.				
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).				

CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
 CC ACYLTRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF375789; AAK54809.1; ALT\_INIT.  
 DR EMBL: ALJ36587; CAB66522.1; ALT\_INIT.  
 DR EMBL: AK002072; BAA92069.1; -.  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KM Phospholipid biosynthesis; Transferase; Acyltransferase;  
 KM Transmembrane.  
 FT TRANSMEM 7 29 POTENTIAL.  
 FT TRANSMEM 44 66 POTENTIAL.  
 FT TRANSMEM 334 351 POTENTIAL.  
 FT CONFLICT 145 145 L -> V (IN REF. 2).  
 SO SEQUENCE 353 AA; 40813 MW; A05B1FA246CE1B64 CRC64;

Query Match 100.0%; Score 1203; DB 1; Length 353;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-104;  
 Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTPAPYVLAMGWRLSAPLPAFYQALDDRLCYQSWVLPFFENYTG 60  
 DB 1 MRYLLPSVLLGTPAPYVLAMGWRLSAPLPAFYQALDDRLCYQSWVLPFFENYTG 60  
 QY 61 VOILLXGDLPRKKNENIYLLANHOSTVDVIADILAIROMALGHVRYLKGKMLPLXGC 120  
 DB 61 VOILLXGDLPRKKNENIYLLANHOSTVDVIADILAIROMALGHVRYLKGKMLPLXGC 120  
 QY 121 YFAOHGCIYVRSKAKFENEKRNKLOSVDAGTPMYLVIFPESTRYNEPQTKVLSAQAF 180  
 DB 121 YFAOHGCIYVRSKAKFENEKRNKLOSVDAGTPMYLVIFPESTRYNEPQTKVLSAQAF 180  
 QY 181 AAORGGLAVLKHVLTLPRIKATHVAFDCKKNYLDLIDYTVVYEGKDDG 228  
 DB 181 AAORGGLAVLKHVLTLPRIKATHVAFDCKKNYLDLIDYTVVYEGKDDG 228

RESULT 2  
 PLCD\_MOUSE ID PLCD\_MOUSE STANDARD; PRT; 354 AA.

AC 09D1E8;  
 DT 01-MAR-2002 (Rel. 41, Created)  
 DT 01-MAR-2002 (Rel. 41, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE 1-acyl-sn-glycerol-3-phosphate acyltransferase epsilon (EC 2.3.1.51)  
 DE (1-AGP acyltransferase 5) (1-AGPAT 5) (Lysophosphatidic acid  
 DE acyltransferase-epsilon) (LPAAT-epsilon) (1-acylglycerol-3-phosphate  
 DE O-acyltransferase 5).  
 GN AGPAT5.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 ON NCBI\_TaxID=10090;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Embryo;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawal J., Shingawa A., Shidata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,  
 RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,  
 RA Fledschmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,

RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmink L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 CC -1- FUNCTION: CONVERTS LYSOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC  
 CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION (BY  
 CC SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate -  
 CC CoA + 1,2-diacyl-sn-glycerol 3-phosphate.  
 CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
 CC ACYLTRANSFERASE FAMILY.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AK003649; BAB22915.1; -.  
 DR MGD: MGI:1915880; Agpat5.  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KM Phospholipid biosynthesis; Transferase; Acyltransferase;  
 KM Transmembrane.  
 FT TRANSMEM 7 29 POTENTIAL.  
 FT TRANSMEM 44 66 POTENTIAL.  
 FT TRANSMEM 335 352 POTENTIAL.  
 SO SEQUENCE 354 AA; 40943 MW; 447BEF924B91E800 CRC64;

Query Match 86.5%; Score 1040; DB 1; Length 354;  
 Best Local Similarity 85.9%; Pred. No. 4.3e-89;  
 Matches 195; Conservative 19; Mismatches 13; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTPAPYVLAMGWRLSAPLPAFYQALDDRLCYQSWVLPFFENYTG 60  
 DB 1 MRYLLPSVLLGTPAPYVLAMGWRLSAPLPAFYQALDDRLCYQSWVLPFFENYTG 60  
 QY 61 VOILLXGDLPRKKNENIYLLANHOSTVDVIADILAIROMALGHVRYLKGKMLPLXGC 120  
 DB 61 VOILLXGDLPRKKNENIYLLANHOSTVDVIADILAIROMALGHVRYLKGKMLPLXGC 120  
 QY 121 YFAOHGCIYVRSKAKFENEKRNKLOSVDAGTPMYLVIFPESTRYNEPQTKVLSAQAF 180  
 DB 121 YFAOHGCIYVRSKAKFENEKRNKLOSVDAGTPMYLVIFPESTRYNEPQTKVLSAQAF 180  
 QY 181 AAORGGLAVLKHVLTLPRIKATHVAFDCKKNYLDLIDYTVVYEGKDDG 227  
 DB 181 AAORGGLAVLKHVLTLPRIKATHVAFDCKKNYLDLIDYTVVYEGKDDG 227

RESULT 3  
 PLCD\_RAT ID PLCD\_RAT STANDARD; PRT; 378 AA.

AC 0924S1;  
 DT 01-MAR-2002 (Rel. 41, Created)  
 DT 01-MAR-2002 (Rel. 41, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE 1-acyl-sn-glycerol-3-phosphate acyltransferase delta (EC 2.3.1.51) (1-  
 DE AGP acyltransferase 4) (1-AGPAT 4) (Lysophosphatidic acid

DE acyltransferase-delta) (LPAAT-delta) (1-acylglycerol-3-phosphate O-  
acetyltransferase 4).  
GN AGPAT4.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Li W., Suzuki T.:  
RT "Rattus norvegicus mRNA for lysophosphatidic acid  
acetyltransferase-delta, complete cds."  
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: CONVERTS LYSPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC  
CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION (BY  
CC SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate =  
CC CoA + 1,2-diacyl-sn-glycerol 3-phosphate.  
CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).  
CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
CC ACYLTRANSFERASE FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL: AB067572; BAB62290.1; -  
KM Phospholipid biosynthesis; Transferase; Acyltransferase;  
KW Transmembrane  
FT TRANSMEM 11 31 POTENTIAL.  
FT TRANSMEM 125 145 POTENTIAL.  
FT TRANSMEM 307 327 POTENTIAL.  
FT TRANSMEM 338 358 POTENTIAL.  
SQ SEQUENCE 378 AA; 43794 MW; 389AA01B7327AE2B CRC64;

Query Match 16.4%; Score 197.5; DB 1; Length 378;  
Best Local Similarity 27.4%; Pred. No. 6.2e-11;  
Matches 55; Conservative 46; Mismatches 85; Indels 15; Gaps 7;  
QY 34 REFQALDDRL-YGVYSMWLEFFENTGVQILLYGDLPR-----KKNITYLANHSTYVD 87  
DB 44 QLEFKINARLCYCV-SSQVLMLEWMSGTECTITYTD-PRASPHYGKEMALIVLNHKEID 101  
QY 88 WIYADILAIRONALGHVRYVLEKGLKMLPLYG--CYFAOHGGIYVKSAREKEMRNKL 145  
DB 102 FLGNSLAERLGLGSKYLAKELAYVPLIGMMYFVEN--IFCRKMEQDQOTYAKSL 159  
QY 146 QSYVDAGTPMYLIFPEGTRYNPEQTVLSASQAFQAORGLAVLKHYLPRIKATHVAFD 205  
DB 160 LHLRDYPERKYLIFHCSEGRFTEKKHDI--SMQVQAQKGLPRLKHHLLPRTGFAITYK 216  
QY 206 CMKNYDAITYDYTVYVEGKDD 226  
DB 217 CLRDVVPAYVDCTLNFRRNEN 237

RESULT 4  
ID PLCC\_HUMAN STANDARD; PRT; 378 AA.  
AC Q9NR25;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE 1-acyl-sn-glycerol-3-phosphate acyltransferase delta (EC 2.3.1.51) (1-  
DE AGP acyltransferase 4) (1-AGPAT 4) (lysophosphatidic acid  
DE acyltransferase-delta) (LPAAT-delta) (1-acylglycerol-3-phosphate O-  
DE acyltransferase 4).

GN AGPAT4.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Leung D.W.;  
RT "Structure and functions of lysophosphatidic acid acyltransferases";  
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: CONVERTS LYSPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC  
CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION (BY  
CC SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate =  
CC CoA + 1,2-diacyl-sn-glycerol 3-phosphate.  
CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).  
CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
CC ACYLTRANSFERASE FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL: AF156776; AAF80338.1; -  
DR InterPro: IPR002123; Acyltransferase.  
DR Pfam: PF01553; Acyltransferase; 1.  
KM Phospholipid biosynthesis; Transferase; Acyltransferase;  
KW Transmembrane  
FT TRANSMEM 11 31 POTENTIAL.  
FT TRANSMEM 125 145 POTENTIAL.  
FT TRANSMEM 307 327 POTENTIAL.  
FT TRANSMEM 338 358 POTENTIAL.  
SQ SEQUENCE 378 AA; 44021 MW; 3EF013D196F8C0C5 CRC64;

Query Match 16.3%; Score 196; DB 1; Length 378;  
Best Local Similarity 24.8%; Pred. No. 8.6e-11;  
Matches 57; Conservative 55; Mismatches 96; Indels 22; Gaps 8;  
QY 5 LPSVILLGTAPRY-VLAWGVRLLSFLPARFQALDDRL-YGVYSMWLEFFENTGVQ 62  
DB 22 IASGLIINTQTLLTLLMPINKOL-----FRKINCRSLSTCI-SSQVLMLEWMSGTE 72  
QY 63 ILLYGD---LPKNKENIITYLANHSTYVDWIYADILAIRONALGHVRYVLEKGLKMLPLY 118  
DB 73 CTIFTPRALYKKGKEMALIVLNHKEIDPLCGMSLSERGLGSKYLAKRELAVPII 132  
QY 119 G-CYFAOHGGIYVKSAREKEMRNKLOSVDAGTPMYLIFPEGTRYNPEQTVLSA 176  
DB 133 GMMWYFLEM--VPCSRRMEQDRKTVATSLQHLNDYDEKYEFLIHCSTGRFTEKKHEI--- 187  
QY 177 SQAFQAORGLAVLKHYLPRIKATHVAFDQMKYLDIAYTVYVYVEGKDD 226  
DB 188 SMQVQAQKGLPRLKHHLLPRTGFAITYVSLRNVAIVYDCTLNFRRNEN 237

RESULT 5  
ID PLCC\_HUMAN STANDARD; PRT; 376 AA.  
AC Q9NR27; Q9NR26;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE 1-acyl-sn-glycerol-3-phosphate acyltransferase gamma (EC 2.3.1.51) (1-  
DE AGP acyltransferase 3) (1-AGPAT 3) (lysophosphatidic acid  
DE acyltransferase-gamma) (LPAAT-gamma) (1-acylglycerol-3-phosphate O-  
DE acyltransferase 3).  
GN AGPAT3.

OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
OX NCBI\_TaxId=9606;  
RN [1]  
RN SEQUENCE FROM N.A. (ISOFORMS GAMMA-1 AND GAMMA-2).  
RA Leung D.W.;  
RT "Structure and functions of lysophosphatidic acid acyltransferases.";  
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RN SEQUENCE FROM N.A.  
RC TISSUE=Fetal liver;  
RA Nagamine K., Kudoh J., Minoshima S., Kawasaki K., Hase T.,  
RA Shimizu N.;  
RT "Isolation of a novel gene encoding 1-acylglycerol-3-phosphate  
O-acyltransferase 3 (AGPAT3) from the human chromosome 21q22.3.";  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RN SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Strausberg R.;  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: CONVERTS LYSPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC  
CC ACID BY INCORPORATING ACYL MOTIETY AT THE 2 POSITION (BY  
CC SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate -  
CC CoA + 1,2-diacyl-sn-glycerol 3-phosphate.  
CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS, GAMMA-1 (SHOWN HERE) AND GAMMA-  
CC 2: ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
CC ACYLTRANSFERASE FAMILY.  
CC  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch).  
CC  
CC  
CC EMBL; AF156774; AAF80336.1; -  
CC EMBL; AF156775; AAF80337.1; -  
CC EMBL; AB040138; BAB18943.1; -  
CC EMBL; BC011971; AAH11971.1; -  
CC InterPro: IPR002123; Acyltransferase.  
CC Pfam: PF01553; Acyltransferase; 1.  
KW Phospholipid biosynthesis; transferase; Acyltransferase;  
KW Transmembrane; Alternative splicing.  
FT TRANSMEM 15 39 POTENTIAL.  
FT TRANSMEM 126 144 POTENTIAL.  
FT TRANSMEM 308 329 POTENTIAL.  
FT TRANSMEM 336 352 POTENTIAL.  
FT VAPSPIC 1 62 MISSING (IN ISOFORM GAMMA-2).  
SQ SEQUENCE 376 AA; 43381 MW; C12CDBD7CC363852 CnC64;

Query Match	15.6%	Score 187.5	DB: 1	Length 376
Best Local Similarity	27.0%	Pred. No. 5.2e-10		
Matches	53	Conservative	43	Mismatches 89; Indels 11; Gaps 4;
QY	34	RYQALDRLRLCVYQSMVLFEEFENTGVQILLYGDLPK----	NKENIYLANHOSTVDMI	89
		: :	: : : : : : : : : : : :	
Db	44	QLYRLNCRRLAVSLMSQLMILEMNSCHTECLFTFOATYERGRKREHAYIILNHNEDIFL		103
QY	90	VADILAIQNALGHRVYVLKEGLKMWLPYLG--CYFAQHGGIVYKRSANFENKEMKNKIOS		147
		: :	: : : : : : : : : : : :	
Db	104	CGWTCERFSGVGLSSKVLAKKELLVPLIGMTWYFLER--VFCKRKWEEDRDTVEGLRR		161
QY	148	YVDAGTPMYLVIFPFGSTIRNEPQRTVYLSAQSAAQRCGLAVLKAHVLTPIKATIHVAFDM		207
		: :	: : : : : : : : : : : :	
Db	162	LSDPDEVMVFLCYGCTFTETTERKHNV---SMEVAANKGPIYKYNILPRTKGFTHAVNCL		218
		: :	: : : : : : : : : : : :	

Oy	208	KNYLDIAIDVTVVEG	223
	:	:: :: :: :	
Dd	219	KQTAVAVIDYTLNFRG	234
RESULT	6		
ID	YD18_YEAST	STANDARD;	PRT; 396 AA.
AC	O12185;		
DT	15-JUL-1999 (Rel. 38, Created)		
DT	15-JUL-1999 (Rel. 38, Last sequence update)		
DT	15-JUL-1999 (Rel. 38, Last annotation update)		
DE	Hypothetical 45.9 kDa protein in KCS1-GCV1 intergenic region.		
GN	YDR018C OR Y09335.04C OR PZF316.		
OS	Saccharomyces cerevisiae (Baker's yeast).		
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;		
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycos.		
OX	NCBI_TaxID=4932;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=S288C / AB972;		
RA	Deelman K., Brown D., Hamlyn N., Bowman S., Barrell B.G.,		
RA	Rajandream M.A.;		
RL	Submitted (May-1995) to the EMBL/GenBank/DDBJ databases.		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=97051598; PubMed=8896275;		
RA	Eide L.G., Sander C., Prydz H.;		
RT	"Sequencing and analysis of a 35.4 kb region on the right arm of		
RT	chromosome IV from Saccharomyces cerevisiae reveal 23 open reading		
RT	frames.";		
RL	Yeast 12:1085-1090(1996).		
CC	-I SUBCELLULAR LOCATION: Integral membrane protein (Potential).		
CC	-I SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE		
CC	ACTINTRANSFERASE FAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>		
CC	or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).		
CC	-----		
DR	EMBL; Z74314; CAA98838.1; -;		
DR	EMBL; X95966; CAA65210.1; -;		
DR	EMBL; Z49770; CAA89843.1; -;		
DR	SGD; S0002425; YDR018C.		
DR	InterPro; IPR002123; Acyltransferase.		
PI	Pfam; PF01553; Acyltransferase; 1.		
KW	Hypothetical protein; Phospholipid biosynthesis; Transferase;		
KW	Acyltransferase; Transmembrane.		
FT	TRANSMEM 27 47	POTENTIAL.	
FT	TRANSMEM 69 89	POTENTIAL.	
FT	TRANSMEM 123 143	POTENTIAL.	
FT	TRANSMEM 372 392	POTENTIAL.	
SQ	SEQUENCE 396 AA; 45938 MW; 9F55AB56C8FD44DD CRC64;		

	Query Match	15.5%	Score 186.5	DB 1	Length 396	
	Best Local Similarity	31.0%	Pred. No. 6.9e-10			
	Matches	53	Conservative	26	Mismatches	67
				Indels	25	Gaps
						4
OY	73 KENIYYLANHGSTVDIMVIADILAIQNALGHVRYRLKEGLKMLPLGYCGFAFHGGIYVKR	132				
	:   :             :   :   :   :   :   :   :   :					
Dd	108 KDRAITITINHOMYADWITYLMLSPVNSUGNVYIIILKALQIIPILGGMNNEKFIFLSR	167				
OY	133 SAKFENEKEMRNKL-----OSYVDAGTPLY-VLPPECTGRNPEQTK	172				
	: : :   :   :   :   :   :   :   :   :   :   :   :					
Dd	168 NMQDERKLTSLSVSMDLNARCKGCLTYRKSCYSTSTNSIAANYINLMPEGTNLIS---	Lk 224				
OY	173 VLSAQAFQAARGTL--AVLKHVLLTPRIKATVAFDCKMKNYIDAIDYDVTVVY	221				



DE Probable 1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51)  
DE (1-AGP acyltransferase) (1-AGPAT) (lysophosphatidic acid  
DE acyltransferase) (LPAAT).  
GN SLC1 OR YDL052C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
RN [1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=94012814; PubMed=8408076;  
RA Nagiec M.M., Wells G.B., Lester R.L., Dickson R.C.;  
RT "A suppressor gene that enables *Saccharomyces cerevisiae* to grow  
RT without making sphingolipids encodes a protein that resembles an  
RT *Escherichia coli* fatty acyltransferase.";  
RL J. Biol. Chem. 268:22156-22163(1993).  
RN [2]  
RN SEQUENCE FROM N.A.  
RA Urestarazu L.A., Andre B., Visers S.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RN SEQUENCE OF 182-303 FROM N.A.  
RA Bloeker H., Brandt P.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: MAY BE AN ACYLTRANSFERASE WITH AN ALTERED SUBSTRATE  
CC SPECIFICITY THAT ENABLES IT TO USE A C-26-COA IN PLACE OF THE  
CC C-16 OR C-18-COAS USED BY THE WILD TYPE PROTEIN.  
CC -1- CATALYTIC ACTIVITY: ACYL-CoA + 1-acyl-sn-glycerol 3-phosphate -  
CC CoA + 1,2-diacyl-sn-glycerol 3-phosphate.  
CC -1- PATHWAY: FIRST STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).  
CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
CC ACYLTRANSFERASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@lsb-sib.ch](mailto:license@lsb-sib.ch)).  
CC -----  
DR EMBL: L13282; AAA16514.1; -;  
DR EMBL: Z74100; CA98614.1; -;  
DR PIR: A48600; A48600.  
DR SGD: S0002210; SLC1.  
DR InterPro: IPR002123; Acyltransferase.  
DR Pfam: PF01553; Acyltransferase; 1.  
KW Phospholipid biosynthesis; Transferase; Acyltransferase;  
KW Transmembrane.  
FT TRANSMEM 16 30 POTENTIAL.  
FT VARIANT 44 44 Q -> L (IN SUPPRESSOR ALLELE SLC1-1).  
SQ SEQUENCE 303 AA; 33887 MW; 36ECB8C2659655EB CRC64;

Query Match 10.0%; Score 120.5; DB 1; Length 303;  
Best Local Similarity 29.7%; Pred. No. 0.00067;  
Matches 44; Conservative 20; Mismatches 61; Indels 23; Gaps 6;

DB 45 CVYOSVLFPEENTGTVOILLGDLPRKNKEIITLANHOSVDIVADILAIROMALGHV 104  
DB 50 CFYHVMKLM-----GLDVKVGGEENLAKKPYIMIANHOSLTDIFM-----LGR 94  
QY 105 -----RYVLEGLKMLPLLYGCFYFAOHGIYKRSKFNKEKMKKLSQSYVDAGTPMYLVI 159  
DB 95 FPPCGTATKAKSLKLYVFLGFMALSGTYFLDRSKRQEAIDTLKGLLENVKK-NKRALMV 153  
QY 160 FPEGTR-YNPEQTVLSASQAF-AAQRG 185  
DB 154 FPEGTRSYSELTMFLPKKGAFFHAAOG 181

RESULT 10

YIHG\_ECOLI  
ID YIHG\_ECOLI STANDARD; PRT; 310 AA.  
AC P32129;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein yihg.  
GN YIHG OR B3862.  
OS *Escherichia coli*.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC *Escherichia*.  
OX NCBI\_TaxID=562;  
RN [1]  
RN SEQUENCE FROM N.A.  
RX STRAIN-K12 / MG1655;  
RL MEDLINE=93347969; PubMed=8346018;  
RA Plunkett G., III, Burland V.D., Daniels D.L., Blattner F.R.;  
RT "Analysis of the *Escherichia coli* genome. III. DNA sequence of the  
RT region from 87.2 to 89.2 minutes.";  
RL Nucleic Acids Res. 21:3391-3398(1993).  
CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
CC ACYLTRANSFERASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@lsb-sib.ch](mailto:license@lsb-sib.ch)).  
CC -----  
DR EMBL: L19201; AAB0297.1; -;  
DR EMBL: AE000461; AAC76860.1; -;  
DR PIR: S40808; S40808.  
DR Ecogen: EG1833; Yihg.  
DR InterPro: IPR002123; Acyltransferase.  
DR Pfam: PF01553; Acyltransferase; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 310 AA; 36289 MW; 9FB3F52EB0B186E CRC64;

Query Match 10.0%; Score 120.5; DB 1; Length 310;  
Best Local Similarity 21.4%; Pred. No. 0.00067;  
Matches 50; Conservative 47; Mismatches 86; Indels 51; Gaps 10;

DB 1 MRYLLPSVLLGTAPTIVLAWGVRLSAFLPAFYQALDRLCYVQSVLFFENYTG 60  
DB 43 VKLLLPVPV-----IMRKYSRP-----CPFMKWCQEGALVLLHLN-PH 80  
QY 61 VOILLGDLPRKNEN-IITLANHOSVDIVADILAIROMALGHV--RYVLEGLKMLP 116  
DB 81 LQMEVHGLEGSKKNWLLICNHRSMADIVLCLVLFK-----HIPMKYFLKQOLAWP 135  
QY 117 LYG--C-----YFAOHGIYKRSKFNKEKMKKLSQSYVDAGT-PMYLVIPPEGRYV 167  
DB 136 FLGLACSLMDPEFKKRSRATVLLRHPERKGDVETTRSEKFRLLHPTTIVNEVSRT 195  
QY 168 PEGTRVLSASQAFAGRLAVLKHLVPRIKATVAFDCKKNVLDIAIVDTYVY 221  
DB 196 QEK-----HQQHTSTFQNLPRPKAAGTAMALNVLGKQFDKLLNTLYC 238

RESULT 11  
PLSC\_LIMDO STANDARD; PRT; 281 AA.  
ID PLSC\_LIMDO  
AC 042870: 040120:  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE 1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51) (1-AGP  
DE acyltransferase) (1-AGPAT) (lysophosphatidic acid acyltransferase)  
DE (LPAAT).  
GN PLSC.



OS Limnathes douglasii (Douglas's meadowfoam).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Limnathaceae; Limnathes.  
OX NCBI\_TaxID=28973;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96028122; PubMed=7588719;  
RA Hanke C., Wolter F.P., Coleman J., Petersek G., Frentzen M.;  
RT "A plant acyltransferase involved in triacylglycerol biosynthesis  
RT complements an Escherichia coli sn-1-acylglycerol-3-phosphate  
RT acyltransferase mutant.";  
RL Eur. J. Biochem. 232:806-810(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96046746; PubMed=7579178;  
RA Brown A.P., Brough C.L., Kroon J., Slabas A.R.;  
RT "Identification of a cDNA that encodes a 1-acyl-sn-glycerol-3-  
RT phosphate acyltransferase from Limnathes douglasii.";  
RL Plant Mol. Biol. 29:267-278(1995).  
CC -1- FUNCTION: CONVERTS LYSPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC  
CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION. THIS ENZYME  
CC USES ERUCOYL-COA AS AN ACYL DONOR.  
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate =  
CC CoA + 1,2-diacyl-sn-glycerol 3-phosphate.  
CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).  
CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
CC ACYLTRANSFERASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: X83266; CAA58239.1; -  
DR EMBL: Z46836; CAA86877.1; -  
DR InterPro: IPR002123; Acyltransferase.  
DR Pfam: PF01553; Acyltransferase; 1.  
KW Phospholipid biosynthesis; Transferase; Acyltransferase;  
KW Transmembrane.  
FT TRANSMEM 40 60 POTENTIAL.  
FT TRANSMEM 71 91 POTENTIAL.  
FT TRANSMEM 110 130 POTENTIAL.  
FT CONFLICT 46 46 I -> V (IN REF. 2).  
FT CONFLICT 188 188 R -> G (IN REF. 2).  
FT CONFLICT 262 262 V -> I (IN REF. 2).  
FT CONFLICT 281 281 N -> K (IN REF. 2).  
SQ SEQUENCE 281 AA; 31716 MW; 9C880BD9E492EE2A CRC64;  
  
Query Match 9.9%; Score 119; DB 1; Length 281;  
Best Local Similarity 23.8%; Pred. No. 0.00082;  
Matches 65; Conservative 40; Mismatches 92; Indels 76; Gaps 14;  
  
QY 2 RLPLPSV-----VLGTAPYV-----LAWGVRLSLAPLRFYQAL 39  
DB 13 ROLKPAVATADDDKGVFMVLLSCFRIFVCAIVLITAVAMGL--IMVLLFWPMRIR 70  
QY 40 DRLRY-CVQSWVLFEEFVNTGVQILLYGDLPKKNENIITYLANHOSTVDMIVADIAIQ 98  
DB 71 LGNLTGHITGILVITWY----GIPIKIQGS-EHTKRAIYISNHASPIDAFVYMWLA--- 122  
QY 99 NALGHVRYVLKEGLKMLPLYGCFPAOHGCIYVRSAKFNEKRNKLSQSYVDAGTP--MY 156  
DB 123 -PIGVGVAKKEVI-WYPLIGOLYTLAHHRIDRS---NPAAAIQSMKEVAVRITTEKNS 177  
QY 157 LVIFPGGTRYNEQOTVLSASQAF---AAQGLAVLKHVLPRIKATHVAF----- 204  
DB 178 LIMFPGGTR--SHDGRLLPFKKGFMHIALQSHLPIVPMILT---GTHLAWRKGFPRVRP 231

QY 205 -----DCKKNYLDAIYDVY 219  
DB 232 VPIIVKVLPPINTDDMTVDKIDDIYVAKIHDIYV 264  
  
RESULT 12  
Y773 CAEEL  
ID Y773 CAEEL STANDARD; PRT; 391 AA.  
AC 011087;  
DT 01-NOV-1997 (rel. 35, Created)  
DT 01-NOV-1997 (rel. 35, Last sequence update)  
DT 01-NOV-1997 (rel. 35, Last annotation update)  
DE Hypothetical 44.7 kDa protein C01G10.3 in chromosome X.  
GN C01G10.3.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
OC Rhabditidae; Pelodierinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA Pauley A.;  
RT Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: SOME, TO C. ELEGANS F08G5.2.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: U23526; AAB36850.1; -  
DR WormPep: C01G10.3; CE02449.  
DR InterPro: IPR002123; Acyltransferase.  
KW Hypothetical protein.  
SQ SEQUENCE 391 AA; 44718 MW; FB5C062A0172A8C1 CRC64;  
  
Query Match 9.8%; Score 117.5; DB 1; Length 391;  
Best Local Similarity 25.7%; Pred. No. 0.0017;  
Matches 49; Conservative 29; Mismatches 74; Indels 39; Gaps 11;  
  
QY 58 YTVGVQLIGD--LPKKNENIITYLANHOSTVDMIVADIAIRONALGHR---YVLKEG 111  
DB 98 FVGATVTEGTNLAGYAEKCLLANHGLDHFV---LMQSLNGKSGIRSRMMWYIYNI 154  
QY 112 LKMLPLYGCFPAOHGCIYV-----KRSAPN--EKEMRNKLSQSYVDAGTPMYLVIPPEGT 164  
DB 155 WKITPL-GVMMWISGNFPGVSGKRDVLSRDRHKNSFYK-DYG--WVIMTPEG 209  
QY 165 RYNPEQTKVLSASQAFPAQGLAVLKHVLPRIKATHVAFD-----CKKNY- 210  
DB 210 RL-----YLVKNSGRTEAKKGLKPLDNCVYPRGAAHVAIVDVGPPDDLSMSKCKGGE 265  
QY 211 LDAIYVTVVY 221  
DB 266 IKYIIDATTIGY 276  
  
RESULT 13  
PLSC\_L1MAL  
ID PLSC\_L1MAL STANDARD; PRT; 281 AA.  
AC Q42868;  
DT 15-DEC-1998 (rel. 37, Created)  
DT 15-DEC-1998 (rel. 37, Last sequence update)  
DT 15-DEC-1998 (rel. 37, Last annotation update)  
DE 1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51) (1-AGP  
DE acyltransferase) (1-AGPAT) (Lysophosphatidic acid acyltransferase)  
DE (LPAAT).  
OS Limnathes alba (White meadowfoam).

Query Match	9.6%	Score 116:	DB 1:	Length 281;
Best Local Similarity	24.2%	Pred. No. 0.0016:		
Matches 59;	Conservative 41;	Mismatches 86;	Indels 58;	Gaps 13.
QY	9	VLLGAPYVVLAMGWYRLSAFLPARFYQALDRLY-CVYQSVMLEFFENYTGVOILLYG	67	
DB	46	IVLITA---VAMGL--IMVLLPMPYMRIRIGNLGYHILGGLVIMLV---GIPLEIGG	95	
QY	68	DLPKKEMITILANQSVVDWIVADIIILRQNALGHVRLVLEKGLKMYLPYGCYFRHOGG	127	
DB	96	S-EHTKKAATYISNASHPLDAEFVWMLA---PISTGVAKREVL-WPFLQLQVTLTLANH	149	
QY	128	IYVKSAAFEKEKEMNNKLSQSYVDAGTP--MYLVIPEEGTRVNPEDQTVLASQAQF---AA	182	
DB	150	IRIDRS---NPAATQSMEKAVRVITEKNLSLIMPEGRSG--DGRLLPFKKGFVHLAL	204	
QY	183	QRGLAVLKHVLTPLRKATHVAF-----DCKNLYLDATY	215	
DB	205	QSHRPVYMITL---GTHLAWRKQCTFPRVPRVPIVKKLPRLNTDDWTVDKIDYVKKMTH	260	
QY	216	DVTY 219		
DB	261	DIYV 264		
RESULT 14				
PLSC_MYCPN				
ID PLSC_MYCPN	STANDARD:	PRT:	266 AA.	
AC P75479;				
DT 01-NOV-1997 (Rel. 35, Created)				
DT 01-NOV-1997 (Rel. 35, Last sequence update)				
DT				

Query Match	9.5%	Score 114	DB 1	Length 266
Best Local Similarity	25.5%	Pred. No. 0.0022		
Matches 56	Conservative 42	Mismatches 74	Indels 48	Gaps 13

  

QY	1	MRYL-LPSVYLLGTAPTVFLAMGVWRLLSAFLPAFYQALDD-----RLCYQYSQWLF	53
Db	13	LRFLQLSLVLY--LPVPEVL-----MLISLSIAKNYESIDENPPELRFKVVRLVSLF	65
QY	54	FFENTGYQIILLGDLPRN--KENIITYLANHOSYVDIVADILAIRQNALGHRVYLRKEG	111
Db	66	LY--IKGVKVIIVN--PENVPKKALVLYANHSNLDPLILLKAFQKTEGVPPLTFEIAFIE	121
QY	112	LK--WL-----PLGYCYEFAQHGGIYVKKRSAGFENKEMRKLL--OSYVDAGTPMYLVIPRE	162
Db	122	LQDPLWLFKIMLLIDCVFLD-----RKNLKRMASLEQOQOIIINQSTA--LCVEPPE	169
QY	163	GTRYNPEQTQVLSASQAFAQGRGLAVLEKHLVLPRIKATHV	202
Db	170	GTR-----VLSRQIGEFKSGALKAVANAPRIVPLTIV	202

  

RESULT 15			
RENV_FOMPV			
ID VENV_FOMPV	STANDARD:	PRT:	377 AA.
AC P36316;			
DT 01-JUN-1994 (Rel. 29, Created)			
DT 01-JUN-1994 (Rel. 29, Last sequence update)			
DT 16-OCT-2001 (Rel. 40, Last annotation update)			
DE Major envelope protein (43 kDa protein) (p43k).			
GN Fowlpox virus (FPV).			
OS Fowlpox virus (FPV).			

```
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Avipoxvirus.
OX NCBI_TaxID=10261;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93079881; PubMed=1333124;
RA Calvert J.G., Ogawa R., Yanagida N., Nazerian K.;
RT "Identification and functional analysis of the fowlpox virus homolog
RT of the vaccinia virus p37k major envelope antigen gene.";
RL Virology 191:783-792(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20193820; PubMed=10729156;
RA Afonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT "The genome of fowlpox virus.";
RL J. Virol. 74:3815-3831(2000).
CC -!- FUNCTION: THIS PROTEIN IS THE MAJOR ANTIGEN ON THE ENVELOPE OF
CC EXTRACELLULAR VIRUS.
CC -!- SIMILARITY: TO VACCINIA VIRUS 37 kDa ENVELOPE PROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC
DR EMBL; M88587; AAA43819.1; -
DR EMBL; M88588; AAA47186.1; -
DR EMBL; AF198100; AAF4452.1; -
DR PIR; AA4216; AA4216.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
DR SMART; SM00155; PLDC; 2.
DR Transmembrane; Late protein.
SQ SEQUENCE 377 AA; 43021 MW; 5F466B7C97F9B86 CRC64;
```

Query Match 9.2%; Score 111; DB 1; Length 377;  
Best Local Similarity 20.9%; Pred. No. 0.0065;  
Matches 51; Conservative 53; Mismatches 80; Indels 60; Gaps 13;

```
OY 1 MRYLLPSVVLGAPRYVYVLAAGVWRLLSAFLPARFYQALDDRLX---CY---QSMVLF 53
DB 128 LHFYISASLSLGNALTYITINMGISYENNSLAMDLYFRSLDYKTISSKKCLEFTPMATKYH 187
OY 54 FFENYGVQVILYGDLPKKKEN-----IYLANHQSIVDWIVADIL----- 94
DB 188 FFRNHNIGI---FFSDSPENHVGKRRTEFDLCVHYITDAKSTIDLAVLSLPTKRTKDSI 244
OY 95 ---AIFONALGHVRYVLEKGLKWLPLYGCFYAGHGSIYVRSRK-FNEKEMRNKLQSYVD 150
DB 245 VYWPILIKDAL--IRAVLEKRGVLRVLG--FWKKTDVVISKASIKSLNELGV-----DHID 295
OY 151 AGTPMWLVIFPESTRYNP-EQTKVLSASQAFAGRLAVLKHYLTPRIKATH-----VA 203
DB 296 ISTKYFR--FPVNSKVDDINNSKMKITIDGRYA-----HVTANLDSHFNHHAFFVS 344
OY 204 FDCM 207
DB 345 FNCM 348
```

Search completed: August 28, 2002, 11:17:24  
Job time: 548 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:16:56 ; Search time 65.95 Seconds

(without alignments)  
598.072 Million cell updates/sec

Title: US-09-853-526-70

Perfect score: 1203

Sequence: 1 MRYLPSVVLGTAPRYVLA.....NYLDATDYVVEKDDG 228

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: SPREMBL\_19:\*  
2: sp\_archaea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_human:\*  
6: sp\_invertebrate:\*  
7: sp\_mammal:\*  
8: sp\_mhc:\*  
9: sp\_organelle:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp Vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1203	100.0	364	4 Q9B0G4	Q9B0G4 homo sapien
2	1200	99.8	353	4 Q9N0G2	Q9N0G2 homo sapien
3	1040	86.5	354	11 Q9D1E8	Q9D1E8 mus musculu
4	338	28.1	386	5 Q9SR12	Q9SR12 caenorhabdi
5	312	25.9	373	10 Q9SSH0	Q9SSH0 arabidopsis
6	312	25.9	393	10 Q9C9P8	Q9C9P8 arabidopsis
7	291	24.2	375	10 Q9LHM4	Q9LHM4 arabidopsis
8	269	22.4	374	10 Q41745	Q41745 zea mays (m
9	240	20.0	306	10 Q9SDN3	Q9SDN3 prunus dulc
10	238	19.8	377	10 Q40119	Q40119 limanthes
11	235.5	19.6	390	10 Q9XFM4	Q9XFM4 brassica na
12	213.5	17.7	376	10 Q9SYC8	Q9SYC8 arabidopsis
13	210	17.5	386	5 Q9VVS1	Q9VVS1 drosophila
14	203	16.9	380	5 Q9VVA9	Q9VVA9 drosophila
15	200	16.6	344	5 Q16526	Q16526 caenorhabdi
16	197.5	16.4	378	11 Q924S1	Q924S1 rattus norv

17	178.5	14.8	316	11 Q9DB84	Q9DB84 mus musculu
18	172.5	14.3	314	11 Q9D517	Q9D517 mus musculu
19	171	14.2	311	10 Q93J17	Q93J17 brassica na
20	168.5	14.0	310	10 Q9SVX9	Q9SVX9 arabidopsis
21	164	13.6	399	5 Q23087	Q23087 caenorhabdi
22	161.5	13.4	350	3 Q94361	Q94361 schizosacch
23	142	11.8	359	10 Q9FF57	Q9FF57 arabidopsis
24	141.5	11.8	426	5 Q01882	Q01882 caenorhabdi
25	141.5	11.8	428	5 Q95R03	Q95R03 caenorhabdi
26	141	11.7	295	16 Q91657	Q91657 pseudomonas
27	134	11.1	439	5 Q20800	Q20800 caenorhabdi
28	133.5	11.1	370	11 Q91YX5	Q91YX5 mus musculu
29	116	9.6	281	10 Q9SDQ2	Q9SDQ2 limanthes
30	113	9.4	247	10 Q9X219	Q9X219 thermotoga
31	112.5	9.4	248	2 Q30849	Q30849 salmonella
32	112	9.3	258	16 Q9HW50	Q9HW50 pseudomonas
33	112	9.3	304	16 Q915A6	Q915A6 pseudomonas
34	109.5	9.1	235	2 Q9KX14	Q9KX14 rhodobacter
35	108	9.0	261	16 Q9KPE3	Q9KPE3 vibrio chol
36	107	8.9	241	16 Q9CPE2	Q9CPE2 pasteurella
37	105.5	8.8	249	2 Q9EY25	Q9EY25 xanthomonas
38	105.5	8.8	262	16 Q98G17	Q98G17 rhizobium 1
39	105	8.7	260	16 Q92MG4	Q92MG4 rhizobium m
40	104	8.6	344	10 Q9LXY4	Q9LXY4 brassica na
41	102.5	8.5	291	3 Q9Y7C5	Q9Y7C5 emericella
42	102	8.5	212	10 Q9M0A2	Q9M0A2 arabidopsis
43	101.5	8.4	292	16 Q9K1N9	Q9K1N9 rhizobium m
44	101	8.4	323	16 Q9KAG4	Q9KAG4 bacillus ba
45	99	8.2	243	16 Q9ZD75	Q9ZD75 rickettsia

## ALIGNMENTS

RESULT 1  
ID Q9B0G4 PRELIMINARY; PRT; 364 AA.  
AC Q9B0G4:  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE HYPOTHEICAL 42.1 KDA PROTEIN (LYSOPHOSPHATIDIC ACID  
DE ACYTRANSEFERASE-EPSILON) (EC 2.3.1.51).  
GN DKFZF61C22 OR LPAAT-E.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxId=9606;  
RN [1]  
RP TISSUE=AMYGDAIA;  
RC MEDLINE=21154917; PubMed=11230166;  
RX Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,  
RA Ansoerge W., Boecker M., Bloecker H., Bauersachs S., Blum H.,  
RA Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,  
RA Mewes H.W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,  
RA Wandt R., Korn B., Klein M., Rouska A.,  
RA "Towards a Catalog of Human Genes and Proteins: Sequencing and  
RT Analysis of 500 Novel Complete Protein Coding Human CDNAs";  
RA Genome Res. 11:422-435(2001).  
[2]  
RP SEQUENCE FROM N.A.  
RA Leung D.W.,  
RT "Cloning and expression of LPAAT-epsilon";  
RT Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL136587; CAB65522.1; -  
DR EMBL; AF375789; AAK54809.1; -  
DR InterPro: IPR002123; Acyltransferase.  
DR Pfam: PF01553; Acyltransferase; 1.  
KW Hypothetical protein; transferase; Acyltransferase.  
SQ SEQUENCE 364 AA: 42072 MW: 90A0F87FC7C78081 CRC64;

Query Match 100.0%; Score 1203; DB 4; Length 364;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-110;  
 Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTAFTYVLAMGWRLLSAFLPARFYQALDDRLCYCYQSWLFEFFENYTG 60  
 |||||  
 DB 12 MRYLLPSVLLGTAFTYVLAMGWRLLSAFLPARFYQALDDRLCYCYQSWLFEFFENYTG 71  
 |||||

QY 61 VOILLGDLPRKKNENIYLANHOSYDWMIVADILAIRONALGHVRYVLEKGLMPLPGC 120  
 |||||  
 DB 72 VOILLGDLPRKKNENIYLANHOSYDWMIVADILAIRONALGHVRYVLEKGLMPLPGC 131  
 |||||

QY 121 YFAQHGGIYVKKRSKAKNEKEMRNKLOSVDAGTPMYLVIFPESTRYNPEOTKYLASQAF 180  
 |||||  
 DB 132 YFAQHGGIYVKKRSKAKNEKEMRNKLOSVDAGTPMYLVIFPESTRYNPEOTKYLASQAF 191  
 |||||

QY 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDIYDVTVVYEEKDDGG 228  
 |||||  
 DB 192 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDIYDVTVVYEEKDDGG 239  
 |||||

RESULT 2  
 ID Q9NU02 PRELIMINARY; PRT: 353 AA.

AC Q9NU02;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)  
 DE CDNA FLJ11210 FIS. CLONE PLACE1007954.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-PLACENTA;  
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
 RA Watsutsumi M., Hosokiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
 RA Yamamoto J., Makamatsu A., Nakamura Y., Nagahari K., Masuno Y.,  
 RA Nimomiya K., Iwayanagi T.;  
 RT "NEO human cDNA sequencing project."  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AK002072; BAA92069.1; -  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 SQ SEQUENCE 353 AA; 40799 MW; 0A8A1A42361F1B14 CRC64;

Query Match 99.8%; Score 1200; DB 4; Length 353;  
 Best Local Similarity 99.6%; Pred. No. 5.6e-110;  
 Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTAFTYVLAMGWRLLSAFLPARFYQALDDRLCYCYQSWLFEFFENYTG 60  
 |||||  
 DB 1 MRYLLPSVLLGTAFTYVLAMGWRLLSAFLPARFYQALDDRLCYCYQSWLFEFFENYTG 60  
 |||||

QY 61 VOILLGDLPRKKNENIYLANHOSYDWMIVADILAIRONALGHVRYVLEKGLMPLPGC 120  
 |||||  
 DB 61 VOILLGDLPRKKNENIYLANHOSYDWMIVADILAIRONALGHVRYVLEKGLMPLPGC 120  
 |||||

QY 121 YFAQHGGIYVKKRSKAKNEKEMRNKLOSVDAGTPMYLVIFPESTRYNPEOTKYLASQAF 180  
 |||||  
 DB 121 YFAQHGGIYVKKRSKAKNEKEMRNKLOSVDAGTPMYLVIFPESTRYNPEOTKYLASQAF 180  
 |||||

QY 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDIYDVTVVYEEKDDGG 228  
 |||||  
 DB 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDIYDVTVVYEEKDDGG 228  
 |||||

RESULT 3

Q9D1E8  
 ID Q9D1E8 PRELIMINARY; PRT: 354 AA.  
 AC Q9D1E8;  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)  
 DE 1110013A05RIK PROTEIN.  
 GN 1110013A05RIK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=EMBRYO;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Glassi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Mashio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boilelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,  
 RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohlsuki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 DR EMBL: AK003649; BAB22915.1; -  
 DR MGI: 1915880; 1110013A05RIK.  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 SQ SEQUENCE 354 AA; 40943 MW; 447EEF924B91E800 CRC64;

Query Match 86.5%; Score 1040; DB 11; Length 354;  
 Best Local Similarity 85.9%; Pred. No. 3.3e-94;  
 Matches 195; Conservative 19; Mismatches 13; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTAFTYVLAMGWRLLSAFLPARFYQALDDRLCYCYQSWLFEFFENYTG 60  
 |||||  
 DB 1 MRYLLPSVLLGTAFTYVLAMGWRLLSAFLPARFYQALDDRLCYCYQSWLFEFFENYTG 60  
 |||||

QY 61 VOILLGDLPRKKNENIYLANHOSYDWMIVADILAIRONALGHVRYVLEKGLMPLPGC 120  
 |||||  
 DB 61 VOILLGDLPRKKNENIYLANHOSYDWMIVADILAIRONALGHVRYVLEKGLMPLPGC 120  
 |||||

QY 121 YFAQHGGIYVKKRSKAKNEKEMRNKLOSVDAGTPMYLVIFPESTRYNPEOTKYLASQAF 180  
 |||||  
 DB 121 YFAQHGGIYVKKRSKAKNEKEMRNKLOSVDAGTPMYLVIFPESTRYNPEOTKYLASQAF 180  
 |||||

QY 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDIYDVTVVYEEKDDGG 227  
 |||||  
 DB 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDIYDVTVVYEEKDDGG 227  
 |||||

RESULT 4  
 ID Q95R12 PRELIMINARY; PRT: 386 AA.  
 AC Q95R12;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)  
 DE HYPOTHETICAL 44.0 KDA PROTEIN.  
 GN F28B3.9.

OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabdilitidae; Pelodolerinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RX MEDLINE=99069613; PubMed=9851916;  
RA None;  
RT "Genome sequence of the nematode C. elegans: a platform for  
RT investigating biology. The C. elegans Sequencing Consortium.";  
RL Science 282:2012-2018(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA Geisel C., Kramer J., Smith A.;  
RT "The sequence of C. elegans cosmid F28B3.";  
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA Waterston R.;  
RT "Direct Submission."  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF003136; AAK93853.1; -.  
KW Hypothetical protein.  
SEQUENCE 386 AA; 43978 MW; A1E3020C201452A6 CRC64;

	Query Match	28.1%	Score 338;	DB 5:	Length 386;	
	Best Local Similarity	35.5%;	Pred. No. 5.9e-25;			
	Matches 83; Conservative	48;	Mismatches 85;	Indels 18;	Gaps 6	
QY	1	KRYLLP----	SVLLGTAPTVLWANGVRKLSAFLPAREYQALDRRLXCVOSMVFEEF	56		
Dd	12	LRLPIPCSLTSSMPWFASCAIVIGV----	SWIWRHVAQDLDMNLKYSYMRCLFEVE	67		
QY	57	NVTGQIIILGXGLPR--KK-----	ENIIYLANQSYVDYDVADILAIHQNALGH---VVRY	107		
Dd	68	NLSGEYLTHGNEEVVNKTGCPENAAVMISNQSNDWTIPMLLAHRHDQDNEGAFRMY	127			
QY	108	LKEGLKMLPLXYGCYEAQHGIYVKRSAKFENEKEMNNKLQSYDAGTPMTLVTFPESTRYN	167			
Dd	128	VKNSTHLVPMEGWYIFQHGVIYVRNFGEIGAPVLROQLKWLNESDPYPVLLLFESTRNS	187			
QY	168	PEQTIVLASCAFAAQROGLVALVLTIRIKATHYAFDOMKNTDAIDYTVVY	221			
Dd	188	AKKHILLSSNNFELEKSGROPOMONVLCPRSSGGLQIALNLNST-LDAIVDYTVY	240			

RESULT	5		
09SSHO			
ID	09SSHO	PRELIMINARY;	PRT; 373 AA.
AC	09SSHO;		
DT	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)		
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)		
DE	F23A4.2 PROTEIN.		
GN	F23A4.2.		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;		
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsids.		
OX	NCBI_TaxID=3702;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=CV. COLUMBIA.		
RA	Vysotskaia V.S., Schwartz J.R., Yu G., Toriumi M., Ienz C., Liu S.,		
RA	Lee J.M., Li J., Gonzalez A., Liu A., Liu K., Vaysberg M., Sakano H.,		
RA	Chin C., Choi E., Chlou J., Altafi H., Araujo R., Brooks S.,		
RA	Buehler E., Chao Q., Conn L., Conway A.B., Dunn P., Hansen N.,		
RA	Hwang B., Huizar L., Khan S., Kim C., Palm C., Rowley D., Shinn P.,		
RA	Walker M., Davis R.W., Eckert J.R., Federspiel N.A., Theologis A.;		

RT "Arabidopsis thaliana chromosome 1 BAC F25n4 sequence,"  
 RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AC0088263; AAD55275.1; -  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase: 1.  
 SQ SEQUENCE 373 AA; 42383 MW; 98642948D9A709B6 CRC64;

[illegible]

RESULT 6

09C9P8	PRELIMINARY;	PRT;	393 AA.
ID 09C9P8			
AC 09C9P8:			
DT 01-JUN-2001	(TREMBLrel. 17, Created)		
DT 01-JUN-2001	(TREMBLrel. 17, last sequence update)		
DT 01-JUN-2001	(TREMBLrel. 17, last annotation update)		
DE PURATIVE ACYL-CoA:1-ACYLGlycerol-3-Phosphate acyltransferase, 31588-29381.			
GN F5E10.13.			
OS Arabidopsis thaliana (Mouse-ear cress).			
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC eustosids II; Brassicales; Brassicaceae; Arabidopsids.			
OX NCBI_TaxId:3702;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC STRAIN=CV. COLUMBIA;			
RX MEDLINE=21016719; PubMed=11130712;			
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S., White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y., Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W., Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K., Dunn P., Etgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y., Gall J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultzer L., Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E., Kim C.J., Koo H.L., Kremetska I., Kurtz D.B., Kwan A., Lam B., Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P., Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Matzali A., Maltischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.T., Paiz G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D., Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M., Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D., Utterback T., Van Aken S., Vaysberg M., Vysotskaya V.S., Walker M., Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;			
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis thaliana.";			
RL Nature 408:816-820(2000).			
DR EMBL: AC013258; AAC51931.1; -			
DR InterPro: IPR002123; Acyltransferase.			
DR Pfam: PF01553; Acyltransferase; 1.			
SW Transferase; Acyltransferase.			
SW SEQUENCE 393 AA; 44892 MW; FA5949D21EBE9A29 CRC64;			





Query Match	19.8%;	Score 238;	DB 10;	length 377;
Best Local Similarity	30.7%;	Pred. No. 4.1e-15;		
Matches	65;	Conservative	49;	Mismatches 74;
				Indels 24;
				Gaps 7;

DT NCBI\_Q93923, 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE PUTATIVE ACYL-COA:1-ACYLGLYCEROL-3-PHOSPHATE ACYLTRANSFERASE.  
 GN FLM15.12.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Federici N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,  
 RA Alatafi H., Araujo R., Huizar L., Rowley D., Buehler E., Dunn P.,  
 RA Gonzalez A., Kremenetskaia I., Kim C., Lenz C., Li J., Liu S.,  
 RA Lueros S., Schwartz J., Shinn P., Toriumi M., Vysotskaia V.S.,



RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Domes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foslter C., Gabriellian A.E., Garg N.S., Gelbatt W.M., Glasser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclebo J.M.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Massarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003527; AAF49473.1; -  
 DR FlyBase: FBgn0036622; CG4753.  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 SQ SEQUENCE 380 AA; 43482 MW; EA240947F566A841 CRC64;

Query Match 16.9%; Score 203; DB 5; Length 380;  
 Best Local Similarity 27.5%; Pred. No. 1.2e-11;  
 Matches 56; Conservative 47; Mismatches 79; Indels 22; Gaps 5;

QY 25 RLISAPLPARFYQALDRKLCVQSWLFFENYTGQIILLXGDLPRN-----KENIIT 78  
 DB 43 RSLAYVHHSFYCI---LVCV-----AEWAGSKLHVYID-PODECKFFGKEHGL 89

QY 79 LANHSTVDIVADILAIKRONALGHVRYLKEGLKMLPLKCYFAQHGITYVKSAPFNE 138  
 DB 90 LMNHTYEIDMLAMITDKLNGGCTKAYAKKMLRYVPLVGWMMMAEFLEDRNFKDK 149

QY 139 KEMRNKLOSVDAGTPMYLIVPEGRTRYNEQGTQVLSASQAFPAQRLAVLKHVLPRIK 198  
 DB 150 VVIKTQLEKFEFSYDPDPVWLLNAGTRFTPAKHEL--SVKFAERGLPLKLHLLIPRTK 206

QY 199 ATHTVAFDCKNKYLDATYDVTVYE 222  
 DB 207 GFTTSLPTMRGICPAIYDINLAFK 230

RESULT 15  
 ID 016526 PRELIMINARY; PRT; 344 AA.  
 AC 016526;  
 DT 01-JAN-1998 (TREMBlrel. 05, Created)  
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE T05H4.1 PROTEIN.  
 GN T05H4.1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_Taxid=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE=94150718; PubMed=7906398;  
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,

RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,  
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,  
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Riffen L., Roopra A., Saunders D., Showkeen R.,  
 RA Saldon N., Smith A., Sonhammer E., Staden R., Sulston J.,  
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
 RA Watson A., Weinstein L., Wilkinson-Sproat J., Wohldman P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of *C.*  
 RT elegans.";  
 RL Nature 368:32-38(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RP STRAIN-BRISTOL N2;  
 RA Blanchard M.;  
 RT "The sequence of *C. elegans* cosmid T05H4.";  
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
 RP [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Waterston R.;  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF016452; AAB66008.1; -  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 SQ SEQUENCE 344 AA; 40070 MW; ABD6C8F47038BE48 CRC64;

Query Match 16.6%; Score 200; DB 5; Length 344;  
 Best Local Similarity 28.7%; Pred. No. 2e-11;  
 Matches 66; Conservative 35; Mismatches 93; Indels 36; Gaps 10;

QY 10 LIGTA----PYVLANGWRLLSAPLPARFYQALDRKLCVQSWLFFENYTGQIILL 65  
 DB 18 LIGTVFLFPLIDPLAW-----FAP-KLMRCADRLVGFWLPFCSLIEWGVNFRV 68

QY 66 YGDLPRKKNIIYLANHSTVDIVADILAIKRONALGHV-----RYLKEGLKMLPL 117  
 DB 69 TGDILIERDEPAIILMNRHRTKLDWLFs-----WNAIKKMPWLLTTEKISLAPLKKIP- 121

QY 118 YGCFYFAQHGIIYVKSAPF-NEKEMRNKLOSVDAGTPMY-LVIFPEGRTRYNEQGTQVLS 175  
 DB 122 -GAGWMASSGSYFLDRNFENDKPVLEIRIVKYSGSEKKYQIILLFAGTDKGERATRL-- 178

QY 176 ASQAFPAQRLAVLKHVLPRIKATHVAFDCKM--ATLDAIYVTVYEG 223  
 DB 179 -SDAFADKNGLPREYVLAHPRTTGFKFLMELMKKENYIKYVYDITAYSG 227

Search completed: August 28, 2002, 11:16:57  
 Job time: 561 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:14:17 ; Search time 30.82 Seconds

(without alignments)  
279.761 Million cell updates/sec

Title: US-09-853-526-4

Perfect score: 353

Sequence: 1 MRYLLPSVLLGTAFTYVLA.....YVNTWIVGTLGLMTVTKA 353

Scoring table:

GAPOP 60.0, Gapext 60.0

Searched: 231628 seqs, 24425594 residues

Word size: 0

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: Issued Patents\_AA:\*

1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/CTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/Dackfilest1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	353	100.0	353	2	US-08-996-306-4
2	353	100.0	353	4	US-09-338-907-4
3	353	100.0	353	4	US-09-218-207-4
4	353	100.0	364	2	US-08-996-306-5
5	353	100.0	364	4	US-09-338-907-5
6	353	100.0	364	4	US-09-218-207-5
7	230	65.2	291	4	US-09-338-907-127
8	230	65.2	291	4	US-09-218-207-127
9	228	64.6	228	4	US-09-338-907-70
10	228	64.6	228	4	US-09-218-207-70
11	199	56.4	261	4	US-09-338-907-128
12	199	56.4	261	4	US-09-218-207-128
13	185	52.4	185	4	US-09-338-907-135
14	185	52.4	185	4	US-09-218-207-135
15	185	52.4	185	4	US-09-338-907-134
16	185	52.4	185	4	US-09-218-207-134
17	184	52.1	300	4	US-09-338-907-135
18	184	52.1	300	4	US-09-218-207-135
19	154	43.6	182	4	US-09-338-907-133
20	154	43.6	182	4	US-09-218-207-133
21	116	32.9	238	4	US-09-338-907-126
22	116	32.9	238	4	US-09-218-207-126
23	85	24.1	97	4	US-09-338-907-132
24	85	24.1	97	4	US-09-218-207-132
25	62	17.6	66	4	US-09-338-907-131
26	62	17.6	66	4	US-09-218-207-131
27	62	17.6	68	4	US-09-338-907-130

28	62	17.6	68	4	US-09-218-207-130	Sequence 130, App
29	62	17.6	77	4	US-09-338-907-125	Sequence 125, App
30	62	17.6	77	4	US-09-218-207-125	Sequence 125, App
31	62	17.6	90	4	US-09-338-907-129	Sequence 129, App
32	62	17.6	90	4	US-09-218-207-129	Sequence 129, App
33	32	9.1	354	4	US-09-338-907-74	Sequence 74, App
34	32	9.1	354	4	US-09-218-207-74	Sequence 74, App
35	9	2.5	9	2	US-08-996-306-15	Sequence 15, App
36	9	2.5	9	4	US-09-338-907-15	Sequence 15, App
37	9	2.5	9	4	US-09-218-207-15	Sequence 15, App
38	8	2.3	9	2	US-08-996-306-18	Sequence 18, App
39	8	2.3	9	4	US-09-338-907-18	Sequence 18, App
40	8	2.3	9	4	US-09-218-207-18	Sequence 18, App
41	7	2.0	12	1	US-08-224-625-1	Sequence 1, App
42	7	2.0	12	2	US-08-254-404-1	Sequence 1, App
43	7	2.0	12	2	US-08-327-451E-1	Sequence 1, App
44	7	2.0	12	2	US-08-458-109-1	Sequence 1, App
45	7	2.0	12	3	US-08-231-196-1	Sequence 1, App

## ALIGNMENTS

RESULT 1  
US-08-996-306-4  
; Sequence 4, Application US/08996306  
; Patent No. 5945522  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Chumakov, Ilya  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Bougenferet, Lydie  
; TITLE OF INVENTION: Prostate cancer gene  
; NUMBER OF SEQUENCES: 68  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson & Bear  
; STREET: 501 West Broadway  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92101-3505  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Win95  
; SOFTWARE: Word  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/996,306  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israelson, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: GENSET.018A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 353 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: SINGLE  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: PROTEIN  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: potential Transmembrane helix  
; LOCATION: 1..33  
; IDENTIFICATION METHOD: Rao and Argos method  
; FEATURE:  
; NAME/KEY: potential Transmembrane helix  
; LOCATION: 4..20

IDENTIFICATION METHOD: Klein, Kanehisa and Delisi method  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 4..24  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony  
IDENTIFICATION METHOD: and Wall method  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 12  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 50..70  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony and Wall method  
FEATURE:  
NAME/KEY: potential N-glycosylation site  
LOCATION: 57  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 76..96  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony and Wall method  
FEATURE:  
NAME/KEY: potential Tyrosine kinase phosphorylation site  
LOCATION: 78  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Casein kinase II phosphorylation site  
LOCATION: 84  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Leucine zipper pattern  
LOCATION: 94..115  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 119  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 133  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Casein kinase II phosphorylation site  
LOCATION: 147  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 194  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Tyrosine kinase phosphorylation site  
LOCATION: 215  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Tyrosine sulfatation site  
LOCATION: 221  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential cAMP- and cGMP-dependent protein kinase phosphorylation site  
LOCATION: 223  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Casein kinase II phosphorylation site  
LOCATION: 235  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 306  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 310..330

IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony and Wall method  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 319  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 323  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Amidation site  
LOCATION: 329  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 333..353  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony and Wall method  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 341  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 350  
IDENTIFICATION METHOD: prosite match  
US-08-996-306-4  
Query Match 100.0%; Score 353; DB 2; Length 353;  
Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRYLPSVVLGTAFTVLAAGVWRLSAFLPAFYQALDRLCYQSVWLEFFENYTG 60  
DB 1 MRYLPSVVLGTAFTVLAAGVWRLSAFLPAFYQALDRLCYQSVWLEFFENYTG 60  
QY 61 VOILLYDLPKKNKNTIYLANHOSTVDMIVADIAIQNALGHRYVLYKSGKMLPYGC 120  
DB 61 VOILLYDLPKKNKNTIYLANHOSTVDMIVADIAIQNALGHRYVLYKSGKMLPYGC 120  
QY 121 YFAOHGSIYVRSKAKFENKEMRNKLOSYVDGTPMYLVIFPEGRTYRPEOTKYVLSAQAF 180  
DB 121 YFAOHGSIYVRSKAKFENKEMRNKLOSYVDGTPMYLVIFPEGRTYRPEOTKYVLSAQAF 180  
QY 181 AAORGGLAVLKHVLPRIKATIHVAFDCKKNYLDATYDVTVVYEGKDDGQRRSEPTWTEFL 240  
DB 181 AAORGGLAVLKHVLPRIKATIHVAFDCKKNYLDATYDVTVVYEGKDDGQRRSEPTWTEFL 240  
QY 241 CKECPRKIHIDRIDKRDVPEDEHMRMWLHERFEIDKMLIEFYESPDEPRRRKRPFGKS 300  
DB 241 CKECPRKIHIDRIDKRDVPEDEHMRMWLHERFEIDKMLIEFYESPDEPRRRKRPFGKS 300  
QY 301 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVNTWITGLGCLWVITKA 353  
DB 301 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVNTWITGLGCLWVITKA 353  
RESULT 2  
US-09-338-907-4  
Sequence 4, Application US/09338907  
Patent No. 6265546  
GENERAL INFORMATION:  
APPLICANT: Cohen, Daniel  
APPLICANT: Blumenfeld, Marla  
APPLICANT: Ilya, Chumakov  
APPLICANT: Bougueleret, Lydie  
TITLE OF INVENTION: PROSTATE CANCER GENE  
FILE REFERENCE: GENSET.18C1CP  
CURRENT APPLICATION NUMBER: US/09/338,907  
CURRENT FILING DATE: 1999-06-23  
EARLIER APPLICATION NUMBER: 08/996,306  
EARLIER FILING DATE: 1997-12-22  
EARLIER APPLICATION NUMBER: 60/099,658

EARLIER FILING DATE: 1998-09-09  
EARLIER APPLICATION NUMBER: 09/218,207  
EARLIER FILING DATE: 1998-12-22  
NUMBER OF SEQ ID NOS: 578  
SOFTWARE: Patent.pm  
SEQ ID NO 4  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 1..33  
OTHER INFORMATION: Rao and Argos identification method, potential helix  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 4..20  
OTHER INFORMATION: Klein, Kanehisa and Delisi identification method, potential  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 4..24  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F  
FEATURE:  
NAME/KEY: MYRISTATE  
LOCATION: 12..16  
OTHER INFORMATION: Prosite match  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 50..70  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F  
FEATURE:  
NAME/KEY: CARBOHYD  
LOCATION: 57..59  
OTHER INFORMATION: Prosite match  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 76..96  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 78  
OTHER INFORMATION: potential Tyrosine kinase site, Prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 84  
OTHER INFORMATION: potential caseine kinase II site, Prosite match  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 94..115  
OTHER INFORMATION: potential leucine zipper site, Prosite match  
FEATURE:  
NAME/KEY: MYRISTATE  
LOCATION: 119..123  
OTHER INFORMATION: potential site, Prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 133  
OTHER INFORMATION: potential protein kinase C, Prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 147  
OTHER INFORMATION: potential caseine kinase II site, Prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 194  
OTHER INFORMATION: potential protein kinase C, Prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 215  
OTHER INFORMATION: potential Tyrosine kinase site, Prosite match  
FEATURE:  
NAME/KEY: SULFATATION  
LOCATION: 221  
OTHER INFORMATION: Prosite match

FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 233  
OTHER INFORMATION: potential cAMP and cGMP dependant protein kinase site,  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 235  
OTHER INFORMATION: potential caseine kinase II site, Prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 306  
OTHER INFORMATION: potential protein kinase C, Prosite match  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 310..330  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
FEATURE:  
NAME/KEY: MYRISTATE  
LOCATION: 319..323  
OTHER INFORMATION: Prosite match  
FEATURE:  
NAME/KEY: MYRISTATE  
LOCATION: 323..327  
OTHER INFORMATION: Prosite match  
FEATURE:  
NAME/KEY: AMIDATION  
LOCATION: 329  
OTHER INFORMATION: Prosite match  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 333..353  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
FEATURE:  
NAME/KEY: MYRISTATE  
LOCATION: 341..345  
OTHER INFORMATION: Prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 350  
OTHER INFORMATION: potential protein kinase C, Prosite match  
US-09-338-907-4  
Query Match 100.0%; Score 353; DB 4; Length 353;  
Best Local Similarity 100.0%; Pred No. 0;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRYLPSVVLGTAPTYVILAMGVWRLLSAFLPARFYQALDDRLCYQSYVTFEFPENYTG 60  
DB 1 MRYLPSVVLGTAPTYVILAMGVWRLLSAFLPARFYQALDDRLCYQSYVTFEFPENYTG 60  
QY 61 VOILLYGDLPKKNENIYYLANHOSYVDWIVADILAIRONALGHVRYLKEGLKWLPLYGC 120  
DB 61 VOILLYGDLPKKNENIYYLANHOSYVDWIVADILAIRONALGHVRYLKEGLKWLPLYGC 120  
QY 121 YFAOHGIIYKRSKFNEMKRNKLOSVDAGTPTLVIPPESTRNPEOTKVLASQAF 180  
DB 121 YFAOHGIIYKRSKFNEMKRNKLOSVDAGTPTLVIPPESTRNPEOTKVLASQAF 180  
QY 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVVYBEKDKGOGORRESPTMTEFL 240  
DB 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVVYBEKDKGOGORRESPTMTEFL 240  
QY 241 CKCEPKIHIDRIKDKVDPEEOEHMRRLHEFEIKDKMLIEFYSPDEPRKRRPPGKS 300  
DB 241 CKCEPKIHIDRIKDKVDPEEOEHMRRLHEFEIKDKMLIEFYSPDEPRKRRPPGKS 300  
QY 301 VNSKLSIKTLPRLMLISGLTAGMLMTDAGRKLVTVMWITGLGLMTWITKA 353  
DB 301 VNSKLSIKTLPRLMLISGLTAGMLMTDAGRKLVTVMWITGLGLMTWITKA 353  
RESULT 3

```
US-09-218-207-4
; Sequence 4, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bouqueloret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; FILE REFERENCE: GENSET.018CP1
; CURRENT APPLICATION NUMBER: US/09/218.207
; EARLIER FILING DATE: 1998-12-22
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 4
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: HELIX
; LOCATION: 1..33
; OTHER INFORMATION: Rao and Argos identification method, potential helix
; FEATURE:
; NAME/KEY: HELIX
; LOCATION: 4..20
; OTHER INFORMATION: Klein, Kanehisa and Delisi identification method, potential
; FEATURE:
; NAME/KEY: HELIX
; LOCATION: 4..24
; OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
; FEATURE:
; NAME/KEY: MYRISTATE
; LOCATION: 12..16
; OTHER INFORMATION: Prosite match
; FEATURE:
; NAME/KEY: HELIX
; LOCATION: 50..70
; OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
; FEATURE:
; NAME/KEY: CARBOHYD
; LOCATION: 57..59
; OTHER INFORMATION: Prosite match
; FEATURE:
; NAME/KEY: HELIX
; LOCATION: 76..96
; OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 78
; OTHER INFORMATION: potential Tyrosine kinase site, Prosite match
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 84
; OTHER INFORMATION: potential caseine kinase II site, Prosite match
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 94..115
; OTHER INFORMATION: potential Leucine zipper site, Prosite match
; FEATURE:
; NAME/KEY: MYRISTATE
; LOCATION: 119..123
; OTHER INFORMATION: potential site, Prosite match
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 133
; OTHER INFORMATION: potential protein kinase C, Prosite match
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 147
; OTHER INFORMATION: potential caseine kinase II site, Prosite match
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 194
; OTHER INFORMATION: potential protein kinase C, Prosite match
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 215
; OTHER INFORMATION: potential Tyrosine kinase site, Prosite match
; FEATURE:
; NAME/KEY: SULFATATION
; LOCATION: 221
; OTHER INFORMATION: Prosite match
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 233
; OTHER INFORMATION: potential CAMP and cGMP dependant protein kinase site,
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 235
; OTHER INFORMATION: potential caseine kinase II site, Prosite match
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 306
; OTHER INFORMATION: potential protein kinase C, Prosite match
; FEATURE:
; NAME/KEY: HELIX
; LOCATION: 310..330
; OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
; FEATURE:
; NAME/KEY: MYRISTATE
; LOCATION: 319..323
; OTHER INFORMATION: Prosite match
; FEATURE:
; NAME/KEY: MYRISTATE
; LOCATION: 323..327
; OTHER INFORMATION: Prosite match
; FEATURE:
; NAME/KEY: AMIDATION
; LOCATION: 329
; OTHER INFORMATION: Prosite match
; FEATURE:
; NAME/KEY: HELIX
; LOCATION: 333..353
; OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
; FEATURE:
; NAME/KEY: MYRISTATE
; LOCATION: 341..345
; OTHER INFORMATION: Prosite match
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 350
; OTHER INFORMATION: potential protein kinase C, Prosite match
; OTHER INFORMATION: potential protein kinase C, Prosite match
US-09-218-207-4

Query Match 100.0%; Score 353; DB 4; Length 353;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVVLGTAPTYVYLAAGWRLISAPLPAFYALDRLCYQSVNLEFFENYTG 60
DB 1 MRYLLPSVVLGTAPTYVYLAAGWRLISAPLPAFYALDRLCYQSVNLEFFENYTG 60
QY 61 VOILLYGDLPKKNENIYILNHOSTYVIVADILAIROMNLGHRVYLKXGLKWLPLYGC 120
DB 61 VOILLYGDLPKKNENIYILNHOSTYVIVADILAIROMNLGHRVYLKXGLKWLPLYGC 120
QY 121 YFAQHGSIYVRSKAFNEKEMRNKLSQSYVDAGTPMYLVIFPEGTRVYPEQTKVLSASQAF 180
DB 121 YFAQHGSIYVRSKAFNEKEMRNKLSQSYVDAGTPMYLVIFPEGTRVYPEQTKVLSASQAF 180
QY 181 AAGRLAVLKHVLTPIKATHVAFDCMKNTLDALIVDTVYVEGKDGQGRRESPTMTTEFL 240
```



|||||  
Db 181 AAGRGALVAKHVLTPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDGGGQRESPTMT EFL 240  
QY 241 CKECPRIHHIDRIKDDVPEDEBHRMRLHEHFEIKDKMLIEFYSPDERRRKRPFGKS 300  
|||||  
Db 241 CKECPRIHHIDRIKDDVPEDEBHRMRLHEHFEIKDKMLIEFYSPDERRRKRPFGKS 300  
QY 301 VNSKLSIKKTLP SMLILSGLTAGMLMTDAGRKLVTWNTWYIGTLGCLMTWITKA 353  
|||||  
Db 301 VNSKLSIKKTLP SMLILSGLTAGMLMTDAGRKLVTWNTWYIGTLGCLMTWITKA 353

## RESULT 4

US-08-966-306-5  
; Sequence 5, Application US/08996306  
; Patent No. 5945522  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Chumakov, Ilya  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: Prostate cancer gene  
; NUMBER OF SEQUENCES: 68  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson & Bear  
; STREET: 501 West Broadway  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92101-3505  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Win95  
; SOFTWARE: Word  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/996,306  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Istaelsen, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: GENSET.018A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 364 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: SINGLE  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: PROTEIN  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: potential protein  
; LOCATION: 1..364  
; US-08-966-306-5

Query Match 100.0%; Score 353; DB 2; Length 364;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRYLLPSVVLGTAPTYVLAAGVWRLLSAFLPARFYQALDDRLCYCYQSMTLFFENYTG 60  
|||||  
Db 12 MRYLLPSVVLGTAPTYVLAAGVWRLLSAFLPARFYQALDDRLCYCYQSMTLFFENYTG 71  
QY 61 VOILLYGDLPKKNENIYYLANHOSYDWMIVADILAIRONALGHVRYVLEKGLKMLPLCYG 120  
|||||  
Db 72 VOILLYGDLPKKNENIYYLANHOSYDWMIVADILAIRONALGHVRYVLEKGLKMLPLCYG 131

QY 121 YFAOHGIIYKRSKAKENEMKRNKLOSVDAGTPTMTLVIPPESTRNPEQTKVLSAQAF 180  
|||||  
Db 132 YFAOHGIIYKRSKAKENEMKRNKLOSVDAGTPTMTLVIPPESTRNPEQTKVLSAQAF 191  
QY 181 AAGRGALVAKHVLTPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDGGGQRESPTMT EFL 240  
Db 192 AAGRGALVAKHVLTPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDGGGQRESPTMT EFL 251  
QY 241 CKECPRIHHIDRIKDDVPEDEBHRMRLHEHFEIKDKMLIEFYSPDERRRKRPFGKS 300  
|||||  
Db 252 CKECPRIHHIDRIKDDVPEDEBHRMRLHEHFEIKDKMLIEFYSPDERRRKRPFGKS 311  
QY 301 VNSKLSIKKTLP SMLILSGLTAGMLMTDAGRKLVTWNTWYIGTLGCLMTWITKA 353  
|||||  
Db 312 VNSKLSIKKTLP SMLILSGLTAGMLMTDAGRKLVTWNTWYIGTLGCLMTWITKA 364

## RESULT 5

US-09-338-907-5  
; Sequence 5, Application US/09338907  
; Patent No. 6265546  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Ilya, Chumakov  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: PROSTATE CANCER GENE  
; FILE REFERENCE: GENSET.18CPICP  
; CURRENT APPLICATION NUMBER: US/09/338,907  
; CURRENT FILING DATE: 1999-06-23  
; EARLIER APPLICATION NUMBER: 08/996,306  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: 60/099,658  
; EARLIER FILING DATE: 1998-09-09  
; EARLIER APPLICATION NUMBER: 09/218,207  
; EARLIER FILING DATE: 1998-12-22  
; NUMBER OF SEQ ID NOS: 578  
; SOFTWARE: Patent.pm  
; SEQ ID NO 5  
; LENGTH: 364  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-338-907-5

Query Match 100.0%; Score 353; DB 4; Length 364;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVVLGTAPTYVLAAGVWRLLSAFLPARFYQALDDRLCYCYQSMTLFFENYTG 60  
|||||  
Db 12 MRYLLPSVVLGTAPTYVLAAGVWRLLSAFLPARFYQALDDRLCYCYQSMTLFFENYTG 71  
QY 61 VOILLYGDLPKKNENIYYLANHOSYDWMIVADILAIRONALGHVRYVLEKGLKMLPLCYG 120  
Db 72 VOILLYGDLPKKNENIYYLANHOSYDWMIVADILAIRONALGHVRYVLEKGLKMLPLCYG 131  
QY 121 YFAOHGIIYKRSKAKENEMKRNKLOSVDAGTPTMTLVIPPESTRNPEQTKVLSAQAF 180  
|||||  
Db 132 YFAOHGIIYKRSKAKENEMKRNKLOSVDAGTPTMTLVIPPESTRNPEQTKVLSAQAF 191  
QY 181 AAGRGALVAKHVLTPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDGGGQRESPTMT EFL 240  
|||||  
Db 192 AAGRGALVAKHVLTPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDGGGQRESPTMT EFL 251  
QY 241 CKECPRIHHIDRIKDDVPEDEBHRMRLHEHFEIKDKMLIEFYSPDERRRKRPFGKS 300  
|||||  
Db 252 CKECPRIHHIDRIKDDVPEDEBHRMRLHEHFEIKDKMLIEFYSPDERRRKRPFGKS 311  
QY 301 VNSKLSIKKTLP SMLILSGLTAGMLMTDAGRKLVTWNTWYIGTLGCLMTWITKA 353  
|||||  
Db 312 VNSKLSIKKTLP SMLILSGLTAGMLMTDAGRKLVTWNTWYIGTLGCLMTWITKA 364

```
RESULT 6
US-09-218-207-5
; Sequence 5, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; FILE REFERENCE: GENSET.018CP1
; CURRENT APPLICATION NUMBER: US/09/218,207
; EARLIER FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 5
; LENGTH: 364
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-218-207-5

Query Match
Best Local Similarity 100.0%; Score 353; DB 4; Length 364;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTAPTYVLAAGVWRLSAPFLPARFYQALDRLCYVQSNVLFEEFNTG 60
DB 12 MRYLLPSVLLGTAPTYVLAAGVWRLSAPFLPARFYQALDRLCYVQSNVLFEEFNTG 71
QY 61 VOILLYGDLPKRNKENTIIYLANHSTVMIYADIIAIRONALGHRVYLKEGLKPLPYGC 120
DB 72 VOILLYGDLPKRNKENTIIYLANHSTVMIYADIIAIRONALGHRVYLKEGLKPLPYGC 131
QY 121 YFAHGIGIYVRSKAFENKEMRNKLOSVDAGTPMYLVIEPEGRVYNEPOTKVLASQAF 180
DB 132 YFAHGIGIYVRSKAFENKEMRNKLOSVDAGTPMYLVIEPEGRVYNEPOTKVLASQAF 191
QY 181 AAGRGVLAHLVLTPTKATHVAFDCMKNYLDIYDVTVYEGKDDGQRRSEPTMTFEL 240
DB 192 AAGRGVLAHLVLTPTKATHVAFDCMKNYLDIYDVTVYEGKDDGQRRSEPTMTFEL 251
QY 241 CKCEPKIHIDRIKDVPEDEQEHMRWLHEREIKDKMLIEFYESPDPERRRRPFGKS 300
DB 252 CKCEPKIHIDRIKDVPEDEQEHMRWLHEREIKDKMLIEFYESPDPERRRRPFGKS 311
QY 301 VNSKLSIKTLPMSMLISGLTAGMLMTDAGRKLTVNTWYIGTLGCLMWTIKA 353
DB 312 VNSKLSIKTLPMSMLISGLTAGMLMTDAGRKLTVNTWYIGTLGCLMWTIKA 364

RESULT 7
US-09-338-907-127
; Sequence 127, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CP1CP
; CURRENT APPLICATION NUMBER: US/09/338,907
; EARLIER FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
```

```
; EARLIER APPLICATION NUMBER: 09/218,207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 127
; LENGTH: 291
; TYPE: PR1
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 98..103
; OTHER INFORMATION: Box II
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 149..157
; OTHER INFORMATION: Box III
US-09-338-907-127

Query Match
Best Local Similarity 100.0%; Score 230; DB 4; Length 291;
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 124 QHGIGIYVRSKAFENKEMRNKLOSVDAGTPMYLVIEPEGRVYNEPOTKVLASQAF 183
DB 62 QHGIGIYVRSKAFENKEMRNKLOSVDAGTPMYLVIEPEGRVYNEPOTKVLASQAF 121
QY 184 RGLAVLKHVLTPTKATHVAFDCMKNYLDIYDVTVYEGKDDGQRRSEPTMTFELCKE 243
DB 122 RGLAVLKHVLTPTKATHVAFDCMKNYLDIYDVTVYEGKDDGQRRSEPTMTFELCKE 181
QY 244 CPKIHIDRIKDVPEDEQEHMRWLHEREIKDKMLIEFYESPDPERRRRPFGKS 303
DB 182 CPKIHIDRIKDVPEDEQEHMRWLHEREIKDKMLIEFYESPDPERRRRPFGKS 241
QY 304 KLSIKTLPMSMLISGLTAGMLMTDAGRKLTVNTWYIGTLGCLMWTIKA 353
DB 242 KLSIKTLPMSMLISGLTAGMLMTDAGRKLTVNTWYIGTLGCLMWTIKA 291

RESULT 8
US-09-218-207-127
; Sequence 127, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; FILE REFERENCE: GENSET.018CP1
; CURRENT APPLICATION NUMBER: US/09/218,207
; EARLIER FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 127
; LENGTH: 291
; TYPE: PR1
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 98..103
; OTHER INFORMATION: Box II
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 149..157
; OTHER INFORMATION: Box III
US-09-218-207-127
```

```
Query Match          65.2%; Score 230; DB 4; Length 291;
Best Local Similarity 100.0%; Pred. No. 6e-223;
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 124 OHGGIYKRSKAFNEKEMRNKLOSVDAGTPTMYLVIFPEGSTRYNPEQTVLSASQAFAAQ 183
    |||||||
DB 62 OHGGIYKRSKAFNEKEMRNKLOSVDAGTPTMYLVIFPEGSTRYNPEQTVLSASQAFAAQ 121
OY 184 RCLAVLKHVLPRIKATHAFDCMKNYLDAIDVTYVYEGKDDGGORRRSPMTTEFLCKE 243
    |||||||
DB 122 RCLAVLKHVLPRIKATHAFDCMKNYLDAIDVTYVYEGKDDGGORRRSPMTTEFLCKE 181
OY 244 CPKIHIDRIDKDVPEEEOEMHRMLHERFEIKDKMLIEFESPDPERRRKPPGSKSVNS 303
    |||||||
DB 182 CPKIHIDRIDKDVPEEEOEMHRMLHERFEIKDKMLIEFESPDPERRRKPPGSKSVNS 241
OY 304 KLSIKKTPSMLISGLTAGMLMTDAGRKLVTYNTWITGTLGCLWVTYKA 353
    |||||||
DB 242 KLSIKKTPSMLISGLTAGMLMTDAGRKLVTYNTWITGTLGCLWVTYKA 291
```

```
RESULT 9
US-09-338-907-70
; Sequence 70, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CPICP
; CURRENT APPLICATION NUMBER: US/09/338,907
; EARLIER FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218,207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 70
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-338-907-70
```

```
Query Match          64.6%; Score 228; DB 4; Length 228;
Best Local Similarity 100.0%; Pred. No. 4.9e-221;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRYLLPSVLLGTAPTYYVLANGVWRLLSAFLPARFYQALDDRLCYQOSWVLEFFENYTG 60
    |||||||
DB 1 MRYLLPSVLLGTAPTYYVLANGVWRLLSAFLPARFYQALDDRLCYQOSWVLEFFENYTG 60
OY 61 VOILLYGDLPKKNENIITYLANHOSYDWDIADILAIRONALGHVRYVLEKGLKWLPLYGC 120
    |||||||
DB 61 VOILLYGDLPKKNENIITYLANHOSYDWDIADILAIRONALGHVRYVLEKGLKWLPLYGC 120
OY 121 YFAOHGIGIYKRSKAFNEKEMRNKLOSVDAGTPTMYLVIFPEGSTRYNPEQTVLSASQAF 180
    |||||||
DB 121 YFAOHGIGIYKRSKAFNEKEMRNKLOSVDAGTPTMYLVIFPEGSTRYNPEQTVLSASQAF 180
OY 181 AAORGLAVLKHVLPRIKATHAFDCMKNYLDAIDVTYVYEGKDDGG 228
    |||||||
DB 181 AAORGLAVLKHVLPRIKATHAFDCMKNYLDAIDVTYVYEGKDDGG 228
```

```
RESULT 10
US-09-218-207-70
```

```
; Sequence 70, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; FILE REFERENCE: GENSET.018CPI
; CURRENT APPLICATION NUMBER: US/09/218,207
; EARLIER FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 70
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-218-207-70
```

```
Query Match          64.6%; Score 228; DB 4; Length 228;
Best Local Similarity 100.0%; Pred. No. 4.9e-221;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRYLLPSVLLGTAPTYYVLANGVWRLLSAFLPARFYQALDDRLCYQOSWVLEFFENYTG 60
    |||||||
DB 1 MRYLLPSVLLGTAPTYYVLANGVWRLLSAFLPARFYQALDDRLCYQOSWVLEFFENYTG 60
OY 61 VOILLYGDLPKKNENIITYLANHOSYDWDIADILAIRONALGHVRYVLEKGLKWLPLYGC 120
    |||||||
DB 61 VOILLYGDLPKKNENIITYLANHOSYDWDIADILAIRONALGHVRYVLEKGLKWLPLYGC 120
OY 121 YFAOHGIGIYKRSKAFNEKEMRNKLOSVDAGTPTMYLVIFPEGSTRYNPEQTVLSASQAF 180
    |||||||
DB 121 YFAOHGIGIYKRSKAFNEKEMRNKLOSVDAGTPTMYLVIFPEGSTRYNPEQTVLSASQAF 180
OY 181 AAORGLAVLKHVLPRIKATHAFDCMKNYLDAIDVTYVYEGKDDGG 228
    |||||||
DB 181 AAORGLAVLKHVLPRIKATHAFDCMKNYLDAIDVTYVYEGKDDGG 228
```

```
RESULT 11
US-09-338-907-128
; Sequence 128, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CPICP
; CURRENT APPLICATION NUMBER: US/09/338,907
; EARLIER FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218,207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 128
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: SITE
; LOCATION: 68..73
```

OTHER INFORMATION: Box II  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 119..127  
OTHER INFORMATION: Box III  
US-09-338-907-128

Query Match 56.4%; Score 199; DB 4; Length 261;  
Best Local Similarity 100.0%; Pred. No. 7.2e-192;  
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 MYLVIPEEGTRYNPEQTKVLSASQAFNAQGLAVLKHVLPRIKATHVAFDCMKNYLDAI 214  
|  
DB 63 MYLVIPEEGTRYNPEQTKVLSASQAFNAQGLAVLKHVLPRIKATHVAFDCMKNYLDAI 122  
QY 215 YDVTVVEGDDGQRRSPPTMEFLCKECPKIHIDRIDKKDVPPEQEHMRRLHERF 274  
|  
DB 123 YDVTVVEGDDGQRRSPPTMEFLCKECPKIHIDRIDKKDVPPEQEHMRRLHERF 182  
QY 275 EIKDKMLIEFESPDPERRRKFPCKSVNSKLSIKTLPMSLISGLTAGMLMTDAGRKL 334  
|  
DB 183 EIKDKMLIEFESPDPERRRKFPCKSVNSKLSIKTLPMSLISGLTAGMLMTDAGRKL 242  
QY 335 VNTWIYGTLLGCLMWTIKA 353  
|  
DB 243 VNTWIYGTLLGCLMWTIKA 261

RESULT 12  
US-09-218-207-128  
Sequence 128, Application US/09218207  
Patent No. 6346381

GENERAL INFORMATION:  
APPLICANT: Cohen, Daniel  
APPLICANT: Blumenfeld, Marta  
APPLICANT: Ilya, Chumakov  
APPLICANT: Bougueleret, Lydie  
TITLE OF INVENTION: Prostate cancer gene  
FILE REFERENCE: GENSET.018CPI  
CURRENT APPLICATION NUMBER: US/09/218,207  
EARLIER FILING DATE: 1998-12-22  
EARLIER APPLICATION NUMBER: 08/996,306  
EARLIER FILING DATE: 1997-12-22  
EARLIER APPLICATION NUMBER: 60/099,658  
NUMBER OF SEQ ID NOS: 578  
SOFTWARE: Patent.pm  
SEQ ID NO 128  
LENGTH: 261  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 68..73  
OTHER INFORMATION: Box II  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 119..127  
OTHER INFORMATION: Box III  
US-09-218-207-128

Query Match 56.4%; Score 199; DB 4; Length 261;  
Best Local Similarity 100.0%; Pred. No. 7.2e-192;  
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 MYLVIPEEGTRYNPEQTKVLSASQAFNAQGLAVLKHVLPRIKATHVAFDCMKNYLDAI 214  
|  
DB 63 MYLVIPEEGTRYNPEQTKVLSASQAFNAQGLAVLKHVLPRIKATHVAFDCMKNYLDAI 122  
QY 215 YDVTVVEGDDGQRRSPPTMEFLCKECPKIHIDRIDKKDVPPEQEHMRRLHERF 274  
|  
DB 123 YDVTVVEGDDGQRRSPPTMEFLCKECPKIHIDRIDKKDVPPEQEHMRRLHERF 182

DB 123 YDVTVVEGDDGQRRSPPTMEFLCKECPKIHIDRIDKKDVPPEQEHMRRLHERF 182  
QY 275 EIKDKMLIEFESPDPERRRKFPCKSVNSKLSIKTLPMSLISGLTAGMLMTDAGRKL 334  
|  
DB 183 EIKDKMLIEFESPDPERRRKFPCKSVNSKLSIKTLPMSLISGLTAGMLMTDAGRKL 242  
QY 335 VNTWIYGTLLGCLMWTIKA 353  
|  
DB 243 VNTWIYGTLLGCLMWTIKA 261

RESULT 13  
US-09-338-907-136  
Sequence 136, Application US/09338907  
Patent No. 6265546

GENERAL INFORMATION:  
APPLICANT: Cohen, Daniel  
APPLICANT: Blumenfeld, Marta  
APPLICANT: Ilya, Chumakov  
APPLICANT: Bougueleret, Lydie  
TITLE OF INVENTION: PROSTATE CANCER GENE  
FILE REFERENCE: GENSET.18CPI  
CURRENT APPLICATION NUMBER: US/09/338,907  
EARLIER FILING DATE: 1999-06-23  
EARLIER APPLICATION NUMBER: 08/996,306  
EARLIER FILING DATE: 1997-12-22  
EARLIER APPLICATION NUMBER: 60/099,658  
EARLIER FILING DATE: 1998-09-09  
EARLIER APPLICATION NUMBER: 09/218,207  
NUMBER OF SEQ ID NOS: 578  
SOFTWARE: Patent.pm  
SEQ ID NO 136  
LENGTH: 185  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 81..83  
OTHER INFORMATION: Box I  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 160..165  
OTHER INFORMATION: Box II  
US-09-338-907-136

Query Match 52.4%; Score 185; DB 4; Length 185;  
Best Local Similarity 100.0%; Pred. No. 5.9e-178;  
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVVLGTAPTVLWGVRLSAFLPARFYQALDRLXCVQSWLFFFEYNTG 60  
|  
DB 1 MRYLLPSVVLGTAPTVLWGVRLSAFLPARFYQALDRLXCVQSWLFFFEYNTG 60  
QY 61 VOILLGDLPRKNENIYLANHSTVDVIADILAIROMNLGHVRYLKGKMLPLYGC 120  
|  
DB 61 VOILLGDLPRKNENIYLANHSTVDVIADILAIROMNLGHVRYLKGKMLPLYGC 120  
QY 121 YFAOHGIIYKRSKAFNEKEMRNKLOSVDAGTPMYLVIPEEGTRYNPEQTKVLSASQAF 180  
|  
DB 121 YFAOHGIIYKRSKAFNEKEMRNKLOSVDAGTPMYLVIPEEGTRYNPEQTKVLSASQAF 180  
QY 181 AAORG 185  
|  
DB 181 AAORG 185

RESULT 14  
US-09-218-207-136  
Sequence 136, Application US/09218207  
Patent No. 6346381  
GENERAL INFORMATION:

```
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; FILE REFERENCE: GENSET.018CPI
; CURRENT APPLICATION NUMBER: US/09/218,207
; CURRENT FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 136
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 81..83
; OTHER INFORMATION: Box I
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 160..165
; OTHER INFORMATION: Box II
US-09-218-207-136
```

```
Query Match          52.4%; Score 185; DB 4; Length 185;
Best Local Similarity 100.0%; Pred. No. 5.9e-178;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 MRYLLPSVYLLGTAPTYVLAMGVWRLLSAFLPARFYQALDDRLCYQOSMVLFFPENNTG 60
    |||||||
DB 1 MRYLLPSVYLLGTAPTYVLAMGVWRLLSAFLPARFYQALDDRLCYQOSMVLFFPENNTG 60
OY 61 VOILLYGDLPRKNENIITYLANHOSTVDWIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120
    |||||||
DB 61 VOILLYGDLPRKNENIITYLANHOSTVDWIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120
OY 121 YFAQHGIIYVKRSKAFNEKEMRNKLQSYVDAGTPMVLVIFPEGTRYNPEQTKVLSASQAF 180
    |||||||
DB 121 YFAQHGIIYVKRSKAFNEKEMRNKLQSYVDAGTPMVLVIFPEGTRYNPEQTKVLSASQAF 180
OY 181 AAORG 185
    |||||
DB 181 AAORG 185
```

```
RESULT 15
US-09-338-907-134
; Sequence 134, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CPICP
; CURRENT APPLICATION NUMBER: US/09/338,907
; CURRENT FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218,207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 134
; LENGTH: 315
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 81..83
; OTHER INFORMATION: Box I
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 160..165
; OTHER INFORMATION: Box II
US-09-338-907-134
```

```
Query Match          52.4%; Score 185; DB 4; Length 315;
Best Local Similarity 100.0%; Pred. No. 9.7e-178;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 MRYLLPSVYLLGTAPTYVLAMGVWRLLSAFLPARFYQALDDRLCYQOSMVLFFPENNTG 60
    |||||||
DB 1 MRYLLPSVYLLGTAPTYVLAMGVWRLLSAFLPARFYQALDDRLCYQOSMVLFFPENNTG 60
OY 61 VOILLYGDLPRKNENIITYLANHOSTVDWIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120
    |||||||
DB 61 VOILLYGDLPRKNENIITYLANHOSTVDWIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120
OY 121 YFAQHGIIYVKRSKAFNEKEMRNKLQSYVDAGTPMVLVIFPEGTRYNPEQTKVLSASQAF 180
    |||||||
DB 121 YFAQHGIIYVKRSKAFNEKEMRNKLQSYVDAGTPMVLVIFPEGTRYNPEQTKVLSASQAF 180
OY 181 AAORG 185
    |||||
DB 181 AAORG 185
```

```
Search completed: August 28, 2002, 11:19:21
Job time: 304 sec
```

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:17:02 ; Search time 305.3 Seconds  
(without alignments)  
406.974 Million cell updates/sec

Title: US-09-853-526-4  
Perfect score: 353  
Sequence: 1 MRYLLPSVLLGTAPTYVLA.....YVNTWIVGTLGLMTIKK 353

Scoring table: OLAGO  
Gapop 60.0 , Gapext 60.0

Searched: 3502263 segs, 351980561 residues  
Word size : 0

Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main:\*

- 1: /cgn2\_6/ptodata/2/paa/PCrus.COMB.pep:\*
- 2: /cgn2\_6/ptodata/2/paa/US06.COMB.pep:\*
- 3: /cgn2\_6/ptodata/2/paa/US07.COMB.pep:\*
- 4: /cgn2\_6/ptodata/2/paa/US080.COMB.pep:\*
- 5: /cgn2\_6/ptodata/2/paa/US081.COMB.pep:\*
- 6: /cgn2\_6/ptodata/2/paa/US082.COMB.pep:\*
- 7: /cgn2\_6/ptodata/2/paa/US083.COMB.pep:\*
- 8: /cgn2\_6/ptodata/2/paa/US084.COMB.pep:\*
- 9: /cgn2\_6/ptodata/2/paa/US085.COMB.pep:\*
- 10: /cgn2\_6/ptodata/2/paa/US086.COMB.pep:\*
- 11: /cgn2\_6/ptodata/2/paa/US088.COMB.pep:\*
- 12: /cgn2\_6/ptodata/2/paa/US089.COMB.pep:\*
- 13: /cgn2\_6/ptodata/2/paa/US090.COMB.pep:\*
- 14: /cgn2\_6/ptodata/2/paa/US091.COMB.pep:\*
- 15: /cgn2\_6/ptodata/2/paa/US092.COMB.pep:\*
- 16: /cgn2\_6/ptodata/2/paa/US093.COMB.pep:\*
- 17: /cgn2\_6/ptodata/2/paa/US094.COMB.pep:\*
- 18: /cgn2\_6/ptodata/2/paa/US095.COMB.pep:\*
- 19: /cgn2\_6/ptodata/2/paa/US096.COMB.pep:\*
- 20: /cgn2\_6/ptodata/2/paa/US097.COMB.pep:\*
- 21: /cgn2\_6/ptodata/2/paa/US098.COMB.pep:\*
- 22: /cgn2\_6/ptodata/2/paa/US099.COMB.pep:\*
- 23: /cgn2\_6/ptodata/2/paa/US100.COMB.pep:\*
- 24: /cgn2\_6/ptodata/2/paa/US101.COMB.pep:\*
- 25: /cgn2\_6/ptodata/2/paa/US102.COMB.pep:\*
- 26: /cgn2\_6/ptodata/2/paa/US103.COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	353	100.0	353	1	PCT-US01-01431-59
2	353	100.0	353	1	PCT-US01-01431-79
3	353	100.0	353	1	PCT-US01-11988-1474
4	353	100.0	353	1	PCT-US01-11988-1475
5	353	100.0	353	22	US-09-853-245-1474
6	353	100.0	353	22	US-09-853-245-1475
7	353	100.0	353	22	US-09-853-526-4

8	353	100.0	353	23	US-09-901-484A-4	Sequence 4, Appl1
9	353	100.0	353	23	US-09-915-582-59	Sequence 59, Appl1
10	353	100.0	353	23	US-09-915-582-79	Sequence 79, Appl1
11	353	100.0	353	26	US-60-099-658-4	Sequence 4, Appl1
12	353	100.0	364	18	US-09-436-919-1	Sequence 1, Appl1
13	353	100.0	364	22	US-09-817-910-7	Sequence 7, Appl1
14	353	100.0	364	22	US-09-853-526-5	Sequence 5, Appl1
15	353	100.0	364	23	US-09-901-484A-5	Sequence 5, Appl1
16	353	100.0	364	26	US-60-099-658-5	Sequence 5, Appl1
17	290	82.2	353	18	US-09-488-725A-2736	Sequence 2736, Ap
18	269	76.2	269	1	PCT-US01-01327-47	Sequence 47, Appl1
19	252	71.4	372	18	US-09-488-725A-6308	Sequence 6308, Ap
20	248	70.3	450	26	US-60-245-222-110	Sequence 110, App
21	230	65.2	291	22	US-09-853-526-127	Sequence 127, App
22	230	65.2	291	23	US-09-901-484A-127	Sequence 127, App
23	228	64.6	228	22	US-09-853-526-70	Sequence 70, Appl1
24	228	64.6	228	23	US-09-901-484A-70	Sequence 70, Appl1
25	228	64.6	228	26	US-60-099-658-70	Sequence 128, App
26	199	56.4	261	22	US-09-853-526-128	Sequence 128, App
27	199	56.4	261	23	US-09-901-484A-128	Sequence 128, App
28	185	52.4	185	22	US-09-853-526-136	Sequence 136, App
29	185	52.4	185	23	US-09-901-484A-136	Sequence 136, App
30	185	52.4	315	22	US-09-853-526-134	Sequence 134, App
31	185	52.4	315	23	US-09-901-484A-134	Sequence 134, App
32	184	52.1	300	22	US-09-853-526-135	Sequence 135, App
33	184	52.1	300	23	US-09-901-484A-135	Sequence 135, App
34	154	43.6	182	22	US-09-853-526-133	Sequence 133, App
35	154	43.6	182	23	US-09-901-484A-133	Sequence 133, App
36	153	43.3	176	18	US-09-436-919-5	Sequence 5, Appl1
37	151	42.8	1032	26	US-60-212-413-309	Sequence 309, App
38	151	42.8	1032	26	US-60-229-518-248	Sequence 248, App
39	116	32.9	238	22	US-09-853-526-126	Sequence 126, App
40	116	32.9	238	23	US-09-901-484A-126	Sequence 126, App
41	85	24.1	97	22	US-09-853-526-132	Sequence 132, App
42	85	24.1	97	23	US-09-901-484A-132	Sequence 132, App
43	81	22.9	92	17	US-09-316-123-124	Sequence 124, App
44	81	22.9	92	22	US-09-827-244-124	Sequence 124, App
45	70	19.8	257	1	PCT-US01-08631-31719	Sequence 31719, A

ALIGNMENTS

RESULT 1  
PCT-US01-01431-59  
Sequence 59, Application PC/TUS0101431  
GENERAL INFORMATION:  
APPLICANT: Human Genome Sciences, Inc.  
TITLE OF INVENTION: 17 human secreted proteins  
FILE REFERENCE: PS723PCT  
CURRENT APPLICATION NUMBER: PCT/US01/01431  
CURRENT FILING DATE: 2001-01-17  
PRIOR APPLICATION NUMBER: 60/179,065  
PRIOR FILING DATE: 2000-01-31  
PRIOR APPLICATION NUMBER: 60/180,628  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: 60/231,968  
NUMBER OF SEQ ID NOS: 97  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 59  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
PCT-US01-01431-59

Query Match 100.0%; Score 353; DB 1; Length 353;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 MRYLLPSVLLGTAPTYVLAAGVRLSAFLPARFYQALDDRLCYVQSMVLEFFENYTG 60  
|||||

```
Db 1 MRYLLPSVLLGTAFTYVLAMGVRRLLSAFLPARFYQALDRILCYVQSWVLEFFENYTG 60
QY 61 VOILLYGDLPKKNENIYYLANHOSTVDMIVADILAIROMALGHVRYVLEKGLKMLPLYGC 120
Db 61 VOILLYGDLPKKNENIYYLANHOSTVDMIVADILAIROMALGHVRYVLEKGLKMLPLYGC 120
QY 121 YFAOHGIIYVRSKAKFENEKERNKLOSVDAGTPLYVIEPEGTRYNPEQTKVLSASQAF 180
Db 121 YFAOHGIIYVRSKAKFENEKERNKLOSVDAGTPLYVIEPEGTRYNPEQTKVLSASQAF 180
QY 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDAYDVTVVYEGKDDGQRRSEPTMTTEFL 240
Db 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDAYDVTVVYEGKDDGQRRSEPTMTTEFL 240
QY 241 CKECPKTHIHIDRIKDKVDPEEDEMRRMLHEREFIKDKMLIEFYESPDPERRRRFPFGKS 300
Db 241 CKECPKTHIHIDRIKDKVDPEEDEMRRMLHEREFIKDKMLIEFYESPDPERRRRFPFGKS 300
QY 301 VNSKLSIKTKLPSMLILSGLTAGMLMTDAGRKLTVNTWIYGTLLGCLMWTIKA 353
Db 301 VNSKLSIKTKLPSMLILSGLTAGMLMTDAGRKLTVNTWIYGTLLGCLMWTIKA 353
```

```
RESULT 2
PCT-US01-01431-79
; Sequence 79, Application PC/TUS0101431
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: 17 human secreted proteins
; FILE REFERENCE: PS723PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01431
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179, 065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180, 628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/231, 968
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 79
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-01431-79
```

```
Query Match 100.0%; Score 353; DB 1; Length 353;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MRYLLPSVLLGTAFTYVLAMGVRRLLSAFLPARFYQALDRILCYVQSWVLEFFENYTG 60
Db 1 MRYLLPSVLLGTAFTYVLAMGVRRLLSAFLPARFYQALDRILCYVQSWVLEFFENYTG 60
QY 61 VOILLYGDLPKKNENIYYLANHOSTVDMIVADILAIROMALGHVRYVLEKGLKMLPLYGC 120
Db 61 VOILLYGDLPKKNENIYYLANHOSTVDMIVADILAIROMALGHVRYVLEKGLKMLPLYGC 120
QY 121 YFAOHGIIYVRSKAKFENEKERNKLOSVDAGTPLYVIEPEGTRYNPEQTKVLSASQAF 180
Db 121 YFAOHGIIYVRSKAKFENEKERNKLOSVDAGTPLYVIEPEGTRYNPEQTKVLSASQAF 180
QY 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDAYDVTVVYEGKDDGQRRSEPTMTTEFL 240
Db 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDAYDVTVVYEGKDDGQRRSEPTMTTEFL 240
QY 241 CKECPKTHIHIDRIKDKVDPEEDEMRRMLHEREFIKDKMLIEFYESPDPERRRRFPFGKS 300
Db 241 CKECPKTHIHIDRIKDKVDPEEDEMRRMLHEREFIKDKMLIEFYESPDPERRRRFPFGKS 300
QY 301 VNSKLSIKTKLPSMLILSGLTAGMLMTDAGRKLTVNTWIYGTLLGCLMWTIKA 353
Db 301 VNSKLSIKTKLPSMLILSGLTAGMLMTDAGRKLTVNTWIYGTLLGCLMWTIKA 353
```

```
Db 301 VNSKLSIKTKLPSMLILSGLTAGMLMTDAGRKLTVNTWIYGTLLGCLMWTIKA 353
RESULT 3
PCT-US01-11988-1474
; Sequence 1474, Application PC/TUS0111988
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: PCT/US01/11988
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1474
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-11988-1474
```

```
Query Match 100.0%; Score 353; DB 1; Length 353;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MRYLLPSVLLGTAFTYVLAMGVRRLLSAFLPARFYQALDRILCYVQSWVLEFFENYTG 60
Db 1 MRYLLPSVLLGTAFTYVLAMGVRRLLSAFLPARFYQALDRILCYVQSWVLEFFENYTG 60
QY 61 VOILLYGDLPKKNENIYYLANHOSTVDMIVADILAIROMALGHVRYVLEKGLKMLPLYGC 120
Db 61 VOILLYGDLPKKNENIYYLANHOSTVDMIVADILAIROMALGHVRYVLEKGLKMLPLYGC 120
QY 121 YFAOHGIIYVRSKAKFENEKERNKLOSVDAGTPLYVIEPEGTRYNPEQTKVLSASQAF 180
Db 121 YFAOHGIIYVRSKAKFENEKERNKLOSVDAGTPLYVIEPEGTRYNPEQTKVLSASQAF 180
QY 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDAYDVTVVYEGKDDGQRRSEPTMTTEFL 240
Db 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDAYDVTVVYEGKDDGQRRSEPTMTTEFL 240
QY 241 CKECPKTHIHIDRIKDKVDPEEDEMRRMLHEREFIKDKMLIEFYESPDPERRRRFPFGKS 300
Db 241 CKECPKTHIHIDRIKDKVDPEEDEMRRMLHEREFIKDKMLIEFYESPDPERRRRFPFGKS 300
QY 301 VNSKLSIKTKLPSMLILSGLTAGMLMTDAGRKLTVNTWIYGTLLGCLMWTIKA 353
Db 301 VNSKLSIKTKLPSMLILSGLTAGMLMTDAGRKLTVNTWIYGTLLGCLMWTIKA 353
```

```
RESULT 4
PCT-US01-11988-1475
; Sequence 1475, Application PC/TUS0111988
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: PCT/US01/11988
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
```



; SEQ ID NO 1475  
; LENGTH: 353  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US01-11988-1475

Query Match 100.0%; Score 353; DB 1; Length 353;  
Best local Similarity 100.0%; Pred. No. 0;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLPSVLLGTAPTYVLAMGWRLISAFLPARFYQALDDRLCYVQSMVLEFFENYTG 60  
DB 1 MRYLPSVLLGTAPTYVLAMGWRLISAFLPARFYQALDDRLCYVQSMVLEFFENYTG 60  
QY 61 VQILLGDLPRKKNENIYYLANHOSYDWMIVADILAIQNALGHVRYLAKGKMLPLYGC 120  
DB 61 VQILLGDLPRKKNENIYYLANHOSYDWMIVADILAIQNALGHVRYLAKGKMLPLYGC 120  
QY 121 YFAOHGIIYKRSKAKFEKEMRNKLOSVDAGTPMYLVIPEEGTRVNPEDTKVLSAQAF 180  
DB 121 YFAOHGIIYKRSKAKFEKEMRNKLOSVDAGTPMYLVIPEEGTRVNPEDTKVLSAQAF 180  
QY 181 AAORGIAVLKHLVLPRIKATHVAFDCMKNYLDAIYDVTVYVEGKDGGQRESPTMTFEL 240  
DB 181 AAORGIAVLKHLVLPRIKATHVAFDCMKNYLDAIYDVTVYVEGKDGGQRESPTMTFEL 240  
QY 241 CKECPRKHHIDRIDKKDVEEEOEHMRMWLHERFEIKDKMLIEFYESPDERRKRFPGKS 300  
DB 241 CKECPRKHHIDRIDKKDVEEEOEHMRMWLHERFEIKDKMLIEFYESPDERRKRFPGKS 300  
QY 301 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLIVNTWYIGTLIGCLMWTIKA 353  
DB 301 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLIVNTWYIGTLIGCLMWTIKA 353

RESULT 5  
US-09-833-245-1474  
; Sequence 1474, Application US/09833245  
; GENERAL INFORMATION:  
; APPLICANT: Human Genome Sciences, Inc.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF546PCT  
; CURRENT APPLICATION NUMBER: US/09/833,245  
; PRIOR FILING DATE: 2001-04-12  
; PRIOR APPLICATION NUMBER: 60/229, 358  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR APPLICATION NUMBER: 60/256, 931  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/199, 384  
; PRIOR FILING DATE: 2000-04-25  
; NUMBER OF SEQ ID NOS: 2267  
; SOFTWARE: Patentln Ver. 2.1  
; SEQ ID NO 1474  
; LENGTH: 353  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-833-245-1474

Query Match 100.0%; Score 353; DB 22; Length 353;  
Best local Similarity 100.0%; Pred. No. 0;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRYLPSVLLGTAPTYVLAMGWRLISAFLPARFYQALDDRLCYVQSMVLEFFENYTG 60  
DB 1 MRYLPSVLLGTAPTYVLAMGWRLISAFLPARFYQALDDRLCYVQSMVLEFFENYTG 60  
QY 61 VQILLGDLPRKKNENIYYLANHOSYDWMIVADILAIQNALGHVRYLAKGKMLPLYGC 120  
DB 61 VQILLGDLPRKKNENIYYLANHOSYDWMIVADILAIQNALGHVRYLAKGKMLPLYGC 120  
QY 121 YFAOHGIIYKRSKAKFEKEMRNKLOSVDAGTPMYLVIPEEGTRVNPEDTKVLSAQAF 180

DB 121 YFAOHGIIYKRSKAKFEKEMRNKLOSVDAGTPMYLVIPEEGTRVNPEDTKVLSAQAF 180  
QY 181 AAORGIAVLKHLVLPRIKATHVAFDCMKNYLDAIYDVTVYVEGKDGGQRESPTMTFEL 240  
DB 181 AAORGIAVLKHLVLPRIKATHVAFDCMKNYLDAIYDVTVYVEGKDGGQRESPTMTFEL 240  
QY 241 CKECPRKHHIDRIDKKDVEEEOEHMRMWLHERFEIKDKMLIEFYESPDERRKRFPGKS 300  
DB 241 CKECPRKHHIDRIDKKDVEEEOEHMRMWLHERFEIKDKMLIEFYESPDERRKRFPGKS 300  
QY 301 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLIVNTWYIGTLIGCLMWTIKA 353  
DB 301 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLIVNTWYIGTLIGCLMWTIKA 353

RESULT 6  
US-09-833-245-1475  
; Sequence 1475, Application US/09833245  
; GENERAL INFORMATION:  
; APPLICANT: Human Genome Sciences, Inc.  
; TITLE OF INVENTION: Albumin Fusion Proteins  
; FILE REFERENCE: PF546PCT  
; CURRENT APPLICATION NUMBER: US/09/833,245  
; PRIOR FILING DATE: 2001-04-12  
; PRIOR APPLICATION NUMBER: 60/229, 358  
; PRIOR FILING DATE: 2000-04-12  
; PRIOR APPLICATION NUMBER: 60/256, 931  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: 60/199, 384  
; PRIOR FILING DATE: 2000-04-25  
; NUMBER OF SEQ ID NOS: 2267  
; SOFTWARE: Patentln Ver. 2.1  
; SEQ ID NO 1475  
; LENGTH: 353  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-833-245-1475

Query Match 100.0%; Score 353; DB 22; Length 353;  
Best local Similarity 100.0%; Pred. No. 0;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRYLPSVLLGTAPTYVLAMGWRLISAFLPARFYQALDDRLCYVQSMVLEFFENYTG 60  
DB 1 MRYLPSVLLGTAPTYVLAMGWRLISAFLPARFYQALDDRLCYVQSMVLEFFENYTG 60  
QY 61 VQILLGDLPRKKNENIYYLANHOSYDWMIVADILAIQNALGHVRYLAKGKMLPLYGC 120  
DB 61 VQILLGDLPRKKNENIYYLANHOSYDWMIVADILAIQNALGHVRYLAKGKMLPLYGC 120  
QY 121 YFAOHGIIYKRSKAKFEKEMRNKLOSVDAGTPMYLVIPEEGTRVNPEDTKVLSAQAF 180  
DB 121 YFAOHGIIYKRSKAKFEKEMRNKLOSVDAGTPMYLVIPEEGTRVNPEDTKVLSAQAF 180  
QY 181 AAORGIAVLKHLVLPRIKATHVAFDCMKNYLDAIYDVTVYVEGKDGGQRESPTMTFEL 240  
DB 181 AAORGIAVLKHLVLPRIKATHVAFDCMKNYLDAIYDVTVYVEGKDGGQRESPTMTFEL 240  
QY 241 CKECPRKHHIDRIDKKDVEEEOEHMRMWLHERFEIKDKMLIEFYESPDERRKRFPGKS 300  
DB 241 CKECPRKHHIDRIDKKDVEEEOEHMRMWLHERFEIKDKMLIEFYESPDERRKRFPGKS 300  
QY 301 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLIVNTWYIGTLIGCLMWTIKA 353  
DB 301 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLIVNTWYIGTLIGCLMWTIKA 353  
RESULT 7  
US-09-853-526-4  
; Sequence 4, Application US/09853526  
; GENERAL INFORMATION:

APPLICANT: Cohen, Daniel  
APPLICANT: Blumenfeld, Marta  
APPLICANT: Ilya, Chumakov  
APPLICANT: Bougueleret, Lydie  
TITLE OF INVENTION: PROSTATE CANCER GENE  
FILE REFERENCE: GENSET.18CPICP  
CURRENT APPLICATION NUMBER: US/09/853,526  
CURRENT FILING DATE: 2001-05-11  
PRIORITY APPLICATION NUMBER: 09/338,907  
PRIORITY FILING DATE: 1999-06-23  
PRIORITY APPLICATION NUMBER: 08/996,306  
PRIORITY FILING DATE: 1997-12-22  
PRIORITY APPLICATION NUMBER: 60/099,658  
PRIORITY FILING DATE: 1998-09-09  
PRIORITY APPLICATION NUMBER: 09/218,207  
PRIORITY FILING DATE: 1998-12-22  
NUMBER OF SEQ ID NOS: 578  
SOFTWARE: Patent.pm  
SEQ ID NO 4  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 1..33  
OTHER INFORMATION: Rao and Argos identification method, potential helix  
NAME/KEY: HELIX  
LOCATION: 4..20  
OTHER INFORMATION: Klein, Kanehisa and Delisi identification method, potential  
NAME/KEY: HELIX  
LOCATION: 4..24  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F  
NAME/KEY: MYRISTATE  
LOCATION: 12..16  
OTHER INFORMATION: Prosite match  
NAME/KEY: HELIX  
LOCATION: 50..70  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F  
NAME/KEY: CARBOHYD  
LOCATION: 57..59  
OTHER INFORMATION: Prosite match  
NAME/KEY: HELIX  
LOCATION: 76..96  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 84  
OTHER INFORMATION: potential caseine kinase II site, Prosite match  
NAME/KEY: SITE  
LOCATION: 94..115  
OTHER INFORMATION: potential leucine zipper site, Prosite match  
NAME/KEY: MYRISTATE  
LOCATION: 119..123  
OTHER INFORMATION: potential site, Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 133  
OTHER INFORMATION: potential protein kinase C, Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 147  
OTHER INFORMATION: potential caseine kinase II site, Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 194  
OTHER INFORMATION: potential protein kinase C, Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 215  
OTHER INFORMATION: potential tyrosine kinase site, Prosite match  
NAME/KEY: SULFATATION  
LOCATION: 221

OTHER INFORMATION: Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 233  
OTHER INFORMATION: potential CAMP and cGMP dependant protein kinase site,  
OTHER INFORMATION: match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 235  
OTHER INFORMATION: potential caseine kinase II site, Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 306  
OTHER INFORMATION: potential protein kinase C, Prosite match  
NAME/KEY: HELIX  
LOCATION: 310..330  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
OTHER INFORMATION: helix  
NAME/KEY: MYRISTATE  
LOCATION: 319..323  
OTHER INFORMATION: Prosite match  
NAME/KEY: MYRISTATE  
LOCATION: 323..327  
OTHER INFORMATION: Prosite match  
NAME/KEY: AMIDATION  
LOCATION: 329  
OTHER INFORMATION: Prosite match  
NAME/KEY: HELIX  
LOCATION: 333..353  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
OTHER INFORMATION: helix  
NAME/KEY: MYRISTATE  
LOCATION: 341..345  
OTHER INFORMATION: Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 350  
OTHER INFORMATION: potential protein kinase C, Prosite match  
US-09-853-526-4

Query Match 100.0%; Score 353; DB 22; Length 353;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVVLGTPATYVLAAGVWRLSAFLPAFYALDDRLCYQVQSVLFFPEFYNG 60  
Db 1 MRYLLPSVVLGTPATYVLAAGVWRLSAFLPAFYALDDRLCYQVQSVLFFPEFYNG 60  
QY 61 VOILLYDLPKKNENIYYLANHOSYVDIYADILAIQNALGHVRYVLYKGLKMLPLYGC 120  
Db 61 VOILLYDLPKKNENIYYLANHOSYVDIYADILAIQNALGHVRYVLYKGLKMLPLYGC 120  
QY 121 YFAOHGSIYVRSKAKFNEKEENRKLQSYVDAGTPMYLVIFPEGTRVNPEDTKVLSAQAF 180  
Db 121 YFAOHGSIYVRSKAKFNEKEENRKLQSYVDAGTPMYLVIFPEGTRVNPEDTKVLSAQAF 180  
QY 181 AAORGGLAVLKHVLTPIRKATHVAFDCKKNYLDIYDVTYVEKDDGQGRRESPTMEFL 240  
Db 181 AAORGGLAVLKHVLTPIRKATHVAFDCKKNYLDIYDVTYVEKDDGQGRRESPTMEFL 240  
QY 241 CKECPKIHIDRIDKQDVEEDEMRRWLHEREFEIKDKMLIEFYESPDPERRKRPFGKS 300  
Db 241 CKECPKIHIDRIDKQDVEEDEMRRWLHEREFEIKDKMLIEFYESPDPERRKRPFGKS 300  
QY 301 VNSKLSIKITLPSMLIISGLTAGMLMTDAGRKLIVNTWITGTLGCIWVTIKA 353  
Db 301 VNSKLSIKITLPSMLIISGLTAGMLMTDAGRKLIVNTWITGTLGCIWVTIKA 353

RESULT 8  
US-09-901-484A-4  
; Sequence 4, Application US/09901484A  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Chumakov, Ilya

APPLICANT: Bougueleret, Lydie  
TITLE OF INVENTION: Prostate Cancer Gene  
FILE REFERENCE: GEN-T11XC3D2  
CURRENT APPLICATION NUMBER: US/09/901,484A  
CURRENT FILING DATE: 2001-07-09  
PRIOR APPLICATION NUMBER: US 08/996,306  
PRIOR FILING DATE: 1997-12-22  
PRIOR APPLICATION NUMBER: US 60/099,658  
PRIOR FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: US 09/218,207  
PRIOR FILING DATE: 1998-12-22  
PRIOR APPLICATION NUMBER: US 09/338,907  
PRIOR FILING DATE: 1999-06-23  
PRIOR APPLICATION NUMBER: US 09/853,526  
PRIOR FILING DATE: 2001-05-11  
NUMBER OF SEQ ID NOS: 578  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 4  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: (1)..(33)  
OTHER INFORMATION: Rao and Argos identification method, potential helix  
NAME/KEY: HELIX  
LOCATION: (4)..(20)  
OTHER INFORMATION: Klein, Kanehisa and Delisi identification method, potential helix  
NAME/KEY: HELIX  
LOCATION: (4)..(24)  
OTHER INFORMATION: Eisenberg, Schwarz, Komaromy, Wall identification method,  
OTHER INFORMATION: potential helix  
NAME/KEY: LIPID  
LOCATION: (12)..(16)  
OTHER INFORMATION: MYRISTATE, Prosite match  
NAME/KEY: HELIX  
LOCATION: (50)..(70)  
OTHER INFORMATION: Eisenberg, Schwarz, Komaromy, Wall identification method,  
OTHER INFORMATION: potential helix  
NAME/KEY: CARBOHYD  
LOCATION: (57)..(59)  
OTHER INFORMATION: Prosite match  
NAME/KEY: HELIX  
LOCATION: (76)..(96)  
OTHER INFORMATION: Eisenberg, Schwarz, Komaromy, Wall identification method,  
OTHER INFORMATION: potential helix  
NAME/KEY: MOD\_RES  
LOCATION: (78)..(78)  
OTHER INFORMATION: PHOSPHORYLATION, potential Tyrosine kinase site, Prosite match  
NAME/KEY: MOD\_RES  
LOCATION: (84)..(84)  
OTHER INFORMATION: PHOSPHORYLATION, potential caseine kinase II site, Prosite match  
NAME/KEY: SITE  
LOCATION: (94)..(115)  
OTHER INFORMATION: Potential leucine zipper site, Prosite match  
NAME/KEY: LIPID  
LOCATION: (119)..(123)  
OTHER INFORMATION: MYRISTATE, Prosite match  
NAME/KEY: MOD\_RES  
LOCATION: (133)..(133)  
OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match  
NAME/KEY: MOD\_RES  
LOCATION: (147)..(147)  
OTHER INFORMATION: PHOSPHORYLATION, potential caseine kinase II, Prosite match  
NAME/KEY: MOD\_RES  
LOCATION: (194)..(194)  
OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match  
NAME/KEY: MOD\_RES  
LOCATION: (215)..(215)  
OTHER INFORMATION: PHOSPHORYLATION, potential Tyrosine kinase site, Prosite match  
NAME/KEY: MOD\_RES  
LOCATION: (221)..(221)  
OTHER INFORMATION: SULFATATION, Prosite match

NAME/KEY: MOD\_RES  
LOCATION: (233)..(233)  
OTHER INFORMATION: PHOSPHORYLATION, potential cAMP and cGMP dependant protein klr  
OTHER INFORMATION: site, Prosite match  
NAME/KEY: MOD\_RES  
LOCATION: (235)..(235)  
OTHER INFORMATION: PHOSPHORYLATION, potential caseine kinase II site, Prosite mat  
NAME/KEY: MOD\_RES  
LOCATION: (306)..(306)  
OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match  
NAME/KEY: HELIX  
LOCATION: (310)..(330)  
OTHER INFORMATION: Eisenberg, Schwarz, Komaromy, Wall identification method,  
OTHER INFORMATION: potential helix  
NAME/KEY: LIPID  
LOCATION: (319)..(323)  
OTHER INFORMATION: MYRISTATE, Prosite match  
NAME/KEY: LIPID  
LOCATION: (323)..(327)  
OTHER INFORMATION: MYRISTATE, Prosite match  
NAME/KEY: MOD\_RES  
LOCATION: (329)..(329)  
OTHER INFORMATION: AMIDATION, Prosite match  
NAME/KEY: HELIX  
LOCATION: (333)..(353)  
OTHER INFORMATION: Eisenberg, Schwarz, Komaromy, Wall identification method,  
OTHER INFORMATION: potential helix  
NAME/KEY: LIPID  
LOCATION: (341)..(345)  
OTHER INFORMATION: MYRISTATE, Prosite match  
NAME/KEY: MOD\_RES  
LOCATION: (350)..(350)  
OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match  
US-09-901-484A-4

Query Match 100.0%; Score 353; DB 23; Length 353;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRVLLPSVLLGTAPTYVLMGWRLSAPLPAPFYQALDRLCYVQSWLFEFFENYTG 60  
Db 1 MRYLLPSVLLGTAPTYVLMGWRLSAPLPAPFYQALDRLCYVQSWLFEFFENYTG 60  
QY 61 VQILLYDDEPKRKNIIYILANHOSVDYIADILAIQNALGHVRYLKGKMLPLYGC 120  
Db 61 VQILLYDDEPKRKNIIYILANHOSVDYIADILAIQNALGHVRYLKGKMLPLYGC 120  
QY 121 YFAOHGSIYVRSKAKFNEKEMRNKLSYVDAGTPMYLVIPPEGTRYPPEOTKYLASQAF 180  
Db 121 YFAOHGSIYVRSKAKFNEKEMRNKLSYVDAGTPMYLVIPPEGTRYPPEOTKYLASQAF 180  
QY 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDALYDVTVYEGKDDGQRRSEPTWTEFL 240  
Db 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDALYDVTVYEGKDDGQRRSEPTWTEFL 240  
QY 241 CKCEPKTHIHIDRDKKDVEEEOEHMRMLHEREPEIKDKMLIEYESPDDERRKRRPFGKS 300  
Db 241 CKCEPKTHIHIDRDKKDVEEEOEHMRMLHEREPEIKDKMLIEYESPDDERRKRRPFGKS 300  
QY 301 VNSKLSIKTLPMSMLISGLTAGMLMTDAGRKLVTNNTWIYGTLLGCLMTYIKA 353  
Db 301 VNSKLSIKTLPMSMLISGLTAGMLMTDAGRKLVTNNTWIYGTLLGCLMTYIKA 353

RESULT 9  
US-09-915-582-59  
Sequence 59, Application US/09915582  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 17 Human Secreted Proteins  
FILE REFERENCE: PS723PI  
CURRENT APPLICATION NUMBER: US/09/915,582

;; CURRENT FILING DATE: 2001-07-27  
;; PRIOR APPLICATION NUMBER: PCT/US01/01431  
;; PRIOR FILING DATE: 2001-01-17  
;; PRIOR APPLICATION NUMBER: 60/179,065  
;; PRIOR FILING DATE: 2000-01-31  
;; PRIOR APPLICATION NUMBER: 60/180,628  
;; PRIOR FILING DATE: 2000-02-04  
;; PRIOR APPLICATION NUMBER: 60/231,968  
;; PRIOR FILING DATE: 2000-09-12  
;; NUMBER OF SEQ ID NOS: 97  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 59  
;; LENGTH: 353  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-915-582-59

Query Match 100.0%; Score 353; DB 23; Length 353;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRYLLPSVLLGTAPTYVLAWGVNRLLSAFLPARFYQALDDRLCYVQSWLFFFEENTG 60  
DB 1 MRYLLPSVLLGTAPTYVLAWGVNRLLSAFLPARFYQALDDRLCYVQSWLFFFEENTG 60  
QY 61 VQILLYGDLPKKNENIYYLANHSTVDMIVADILAIRONALGHVRYVLEKGLKMLPLYGC 120  
DB 61 VQILLYGDLPKKNENIYYLANHSTVDMIVADILAIRONALGHVRYVLEKGLKMLPLYGC 120  
QY 121 YFAOHGCIYVRSKAFNKEKRNKLOSVDAGTMYLVIPEEGTRYNPEQTKVLSASQAF 180  
DB 121 YFAOHGCIYVRSKAFNKEKRNKLOSVDAGTMYLVIPEEGTRYNPEQTKVLSASQAF 180  
QY 181 AAORGGLAVLKHLVLPRIKATHVAFDCMKNYLDATYDVTVYVEGKDDGQRRSEPTMTEFL 240  
DB 181 AAORGGLAVLKHLVLPRIKATHVAFDCMKNYLDATYDVTVYVEGKDDGQRRSEPTMTEFL 240  
QY 241 CKECPKIHIDRIDKKDVPEQEHMRRLHEREIKDKMLIEFYESPDEPRRRKRPFGKS 300  
DB 241 CKECPKIHIDRIDKKDVPEQEHMRRLHEREIKDKMLIEFYESPDEPRRRKRPFGKS 300  
QY 301 VNSKLSIKTKLPMLISGLTAGMLMTDAGRKLYVNTWYIGTLGCLMWTIKA 353  
DB 301 VNSKLSIKTKLPMLISGLTAGMLMTDAGRKLYVNTWYIGTLGCLMWTIKA 353

RESULT 10  
US-09-915-582-79  
;; Sequence 79, Application US/09915582  
;; GENERAL INFORMATION:  
;; APPLICANT: Rosen et al.  
;; TITLE OF INVENTION: 17 Human Secreted Proteins  
;; FILE REFERENCE: PS723P1  
;; CURRENT APPLICATION NUMBER: US/09/915,582  
;; PRIOR FILING DATE: 2001-07-27  
;; PRIOR APPLICATION NUMBER: PCT/US01/01431  
;; PRIOR FILING DATE: 2001-01-17  
;; PRIOR APPLICATION NUMBER: 60/179,065  
;; PRIOR FILING DATE: 2000-01-31  
;; PRIOR APPLICATION NUMBER: 60/180,628  
;; PRIOR FILING DATE: 2000-02-04  
;; PRIOR APPLICATION NUMBER: 60/231,968  
;; PRIOR FILING DATE: 2000-09-12  
;; NUMBER OF SEQ ID NOS: 97  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 79  
;; LENGTH: 353  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-915-582-79

Query Match 100.0%; Score 353; DB 23; Length 353;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRYLLPSVLLGTAPTYVLAWGVNRLLSAFLPARFYQALDDRLCYVQSWLFFFEENTG 60  
DB 1 MRYLLPSVLLGTAPTYVLAWGVNRLLSAFLPARFYQALDDRLCYVQSWLFFFEENTG 60  
QY 61 VQILLYGDLPKKNENIYYLANHSTVDMIVADILAIRONALGHVRYVLEKGLKMLPLYGC 120  
DB 61 VQILLYGDLPKKNENIYYLANHSTVDMIVADILAIRONALGHVRYVLEKGLKMLPLYGC 120  
QY 121 YFAOHGCIYVRSKAFNKEKRNKLOSVDAGTMYLVIPEEGTRYNPEQTKVLSASQAF 180  
DB 121 YFAOHGCIYVRSKAFNKEKRNKLOSVDAGTMYLVIPEEGTRYNPEQTKVLSASQAF 180  
QY 181 AAORGGLAVLKHLVLPRIKATHVAFDCMKNYLDATYDVTVYVEGKDDGQRRSEPTMTEFL 240  
DB 181 AAORGGLAVLKHLVLPRIKATHVAFDCMKNYLDATYDVTVYVEGKDDGQRRSEPTMTEFL 240  
QY 241 CKECPKIHIDRIDKKDVPEQEHMRRLHEREIKDKMLIEFYESPDEPRRRKRPFGKS 300  
DB 241 CKECPKIHIDRIDKKDVPEQEHMRRLHEREIKDKMLIEFYESPDEPRRRKRPFGKS 300  
QY 301 VNSKLSIKTKLPMLISGLTAGMLMTDAGRKLYVNTWYIGTLGCLMWTIKA 353  
DB 301 VNSKLSIKTKLPMLISGLTAGMLMTDAGRKLYVNTWYIGTLGCLMWTIKA 353

RESULT 11  
US-60-099-658-4  
;; Sequence 4, Application US/60099658  
;; GENERAL INFORMATION:  
;; APPLICANT: Cohen, Daniel  
;; APPLICANT: Chumakov, Ilya  
;; APPLICANT: Blumenfeld, Marta  
;; APPLICANT: Bougueterec, Lydie  
;; TITLE OF INVENTION: Prostate cancer gene  
;; NUMBER OF SEQUENCES: 99  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Knobbe, Martens, Olson & Bear  
;; STREET: 501 West Broadway  
;; CITY: San Diego  
;; STATE: California  
;; COUNTRY: USA  
;; ZIP: 92101-3505  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: Win95  
;; SOFTWARE: Word  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/60/099,658  
;; FILING DATE:  
;; CLASSIFICATION:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Israelson, Ned A.  
;; REGISTRATION NUMBER: 29,655  
;; REFERENCE/DOCKET NUMBER: GENSET.018APR  
;; TELEPHONE: (619) 235-8550  
;; TELEFAX: (619) 235-0176  
;; INFORMATION FOR SEQ ID NO: 4:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 353 amino acids  
;; TYPE: AMINO ACID  
;; STRANDEDNESS: SINGLE  
;; TOPOLOGY: LINEAR  
;; MOLECULE TYPE: PROTEIN  
;; ORIGINAL SOURCE:  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
NAME/KEY: potential Transmembrane helix

LOCATION: 1..33  
IDENTIFICATION METHOD: Rao and Argos method  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 4..20  
IDENTIFICATION METHOD: Klein, Kanehisa and Delisi method  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 4..24  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komaromy  
IDENTIFICATION METHOD: and Wall method  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 12  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 50..70  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komaromy and Wall method  
FEATURE:  
NAME/KEY: potential N-glycosylation site  
LOCATION: 57  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 76..96  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komaromy and Wall method  
FEATURE:  
NAME/KEY: potential Tyrosine kinase phosphorylation site  
LOCATION: 78  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Casein kinase II phosphorylation site  
LOCATION: 84  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Leucine zipper pattern  
LOCATION: 94..115  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 119  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 133  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Casein kinase II phosphorylation site  
LOCATION: 147  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 194  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Tyrosine kinase phosphorylation site  
LOCATION: 215  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Tyrosine sulfatation site  
LOCATION: 221  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential CAMP- and cGMP-dependent protein kinase phosphorylation site  
LOCATION: 233  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Casein kinase II phosphorylation site  
LOCATION: 235  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site

LOCATION: 306  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 310..330  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komaromy and Wall method  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 319  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 323  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Amidation site  
LOCATION: 329  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 333..353  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komaromy and Wall method  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 341  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 350  
IDENTIFICATION METHOD: prosite match  
US-60-099-658-4

Query Match 100.0%; Score 353; DB 26; Length 353;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRLTLPSSVVLGTPAPTYVLAAGWRLLSAFLPARFYQALDDRLCYQSQSVLFFENYTG 60  
|||||  
DB 1 MRLTLPSSVVLGTPAPTYVLAAGWRLLSAFLPARFYQALDDRLCYQSQSVLFFENYTG 60  
QY 61 VQILLVGDLPKKNENIIVLANHOSYDWMIVADILATRONALGHRVYLEGKMLPLYGC 120  
|||||  
DB 61 VQILLVGDLPKKNENIIVLANHOSYDWMIVADILATRONALGHRVYLEGKMLPLYGC 120  
QY 121 YFAOHGIIYKRSKAKFNEKEMRNKLQSYVDAGTPMYLVIFPESTRYNPEQTKVLSASQAF 180  
|||||  
DB 121 YFAOHGIIYKRSKAKFNEKEMRNKLQSYVDAGTPMYLVIFPESTRYNPEQTKVLSASQAF 180  
QY 181 AAOGLAVLKHVLTPTIKATKTHVAFDCMKNYLDAIVYVYEEGKDDGGQRRSPMTTEPL 240  
|||||  
DB 181 AAOGLAVLKHVLTPTIKATKTHVAFDCMKNYLDAIVYVYEEGKDDGGQRRSPMTTEPL 240  
QY 241 CKCEPKIHIDIRKDDVPEEOEHMRMLHEFEIKDKMLIEFYSPDERKRRPPGKS 300  
|||||  
DB 241 CKCEPKIHIDIRKDDVPEEOEHMRMLHEFEIKDKMLIEFYSPDERKRRPPGKS 300  
QY 301 VNSKLSIKRTLPSPMLILSGITAGMLMTDAGRKLYVTWVIYGTLLGCLMTWYTIKA 353  
|||||  
DB 301 VNSKLSIKRTLPSPMLILSGITAGMLMTDAGRKLYVTWVIYGTLLGCLMTWYTIKA 353

RESULT 12  
US-09-436-919-1  
; Sequence 1, Application US/09436919A  
; GENERAL INFORMATION:  
; APPLICANT: Leung, David W  
; TITLE OF INVENTION: Human Lysophosphatidic Acid Acyltransferase-epsilon  
; FILE REFERENCE: 1801B  
; CURRENT APPLICATION NUMBER: US/09/436,919A  
; CURRENT FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn Ver. 2.0

```
; SEQ ID NO 1
; LENGTH: 364
; TYPE: PRF
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: hLPAATEpsilon
US-09-436-919-1
```

```
Query Match          100.0%; Score 353; DB 18; Length 364;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 MRYLLPSVLLGTAPTYVLAWGVRLLSAFLPAFYQALDRLCYVQSWLFFFEYNTG 60
    |||||
DB 12 MRYLLPSVLLGTAPTYVLAWGVRLLSAFLPAFYQALDRLCYVQSWLFFFEYNTG 71
    |||||
OY 61 VOILLYGDLPKKNENIYYLANHOSTVDMIVADILAIQNALGHRVYLKESGLKMLPLYGC 120
    |||||
DB 72 VOILLYGDLPKKNENIYYLANHOSTVDMIVADILAIQNALGHRVYLKESGLKMLPLYGC 131
    |||||
OY 121 YFAOHGSIYVRSKAFENKEKRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASAOAF 180
    |||||
DB 132 YFAOHGSIYVRSKAFENKEKRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASAOAF 191
    |||||
OY 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDALYDVTVYEGKDGQORRESPTMTTEFL 240
    |||||
DB 192 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDALYDVTVYEGKDGQORRESPTMTTEFL 251
    |||||
OY 241 CKECPKIHIDRIDKDDVEEOEHMRWLHEREIKDKMLIEFYESPDEPRRRRPFSGS 300
    |||||
DB 252 CKECPKIHIDRIDKDDVEEOEHMRWLHEREIKDKMLIEFYESPDEPRRRRPFSGS 311
    |||||
OY 301 VNSKLSIKTLPMSLISGLTAGMLMTDAGRKLVTNTWYIGTLGCLMTWITKA 353
    |||||
DB 312 VNSKLSIKTLPMSLISGLTAGMLMTDAGRKLVTNTWYIGTLGCLMTWITKA 364
    |||||
```

## RESULT 13

```
US-09-817-910-7
; Sequence 7, Application US/09817910
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; APPLICANT: MacBeth, Kyle J.
; APPLICANT: Williamson, Mark
; APPLICANT: Rudolph-Owen, Laura A.
; APPLICANT: Tsai, Fong-Ying
; TITLE OF INVENTION: 46743 and 27417 NOVEL HUMAN
; FILE REFERENCE: ACYLTRANSFERASES AND USES THEREFOR
; CURRENT APPLICATION NUMBER: US/09/817,910
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/192,092
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 364
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-817-910-7
```

```
Query Match          100.0%; Score 353; DB 22; Length 364;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 MRYLLPSVLLGTAPTYVLAWGVRLLSAFLPAFYQALDRLCYVQSWLFFFEYNTG 60
    |||||
DB 12 MRYLLPSVLLGTAPTYVLAWGVRLLSAFLPAFYQALDRLCYVQSWLFFFEYNTG 71
    |||||
OY 61 VOILLYGDLPKKNENIYYLANHOSTVDMIVADILAIQNALGHRVYLKESGLKMLPLYGC 120
    |||||
```

```
DB 72 VOILLYGDLPKKNENIYYLANHOSTVDMIVADILAIQNALGHRVYLKESGLKMLPLYGC 131
    |||||
OY 121 YFAOHGSIYVRSKAFENKEKRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASAOAF 180
    |||||
DB 132 YFAOHGSIYVRSKAFENKEKRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASAOAF 191
    |||||
OY 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDALYDVTVYEGKDGQORRESPTMTTEFL 240
    |||||
DB 192 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDALYDVTVYEGKDGQORRESPTMTTEFL 251
    |||||
OY 241 CKECPKIHIDRIDKDDVEEOEHMRWLHEREIKDKMLIEFYESPDEPRRRRPFSGS 300
    |||||
DB 252 CKECPKIHIDRIDKDDVEEOEHMRWLHEREIKDKMLIEFYESPDEPRRRRPFSGS 311
    |||||
OY 301 VNSKLSIKTLPMSLISGLTAGMLMTDAGRKLVTNTWYIGTLGCLMTWITKA 353
    |||||
DB 312 VNSKLSIKTLPMSLISGLTAGMLMTDAGRKLVTNTWYIGTLGCLMTWITKA 364
    |||||
```

## RESULT 14

```
US-09-853-526-5
; Sequence 5, Application US/09853526
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Cohen, Daniel
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET 18CP1CP
; CURRENT APPLICATION NUMBER: US/09/853,526
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 09/318,907
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: 08/996,306
; PRIOR FILING DATE: 1997-12-22
; PRIOR APPLICATION NUMBER: 60/099,658
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 09/218,207
; PRIOR FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 5
; LENGTH: 364
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-853-526-5
```

```
Query Match          100.0%; Score 353; DB 22; Length 364;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 MRYLLPSVLLGTAPTYVLAWGVRLLSAFLPAFYQALDRLCYVQSWLFFFEYNTG 60
    |||||
DB 12 MRYLLPSVLLGTAPTYVLAWGVRLLSAFLPAFYQALDRLCYVQSWLFFFEYNTG 71
    |||||
OY 61 VOILLYGDLPKKNENIYYLANHOSTVDMIVADILAIQNALGHRVYLKESGLKMLPLYGC 120
    |||||
DB 72 VOILLYGDLPKKNENIYYLANHOSTVDMIVADILAIQNALGHRVYLKESGLKMLPLYGC 131
    |||||
OY 121 YFAOHGSIYVRSKAFENKEKRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASAOAF 180
    |||||
DB 132 YFAOHGSIYVRSKAFENKEKRNKLOSVDAGTPMYLVIFPEGTRYNPEOTKVLASAOAF 191
    |||||
OY 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDALYDVTVYEGKDGQORRESPTMTTEFL 240
    |||||
DB 192 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDALYDVTVYEGKDGQORRESPTMTTEFL 251
    |||||
OY 241 CKECPKIHIDRIDKDDVEEOEHMRWLHEREIKDKMLIEFYESPDEPRRRRPFSGS 300
    |||||
DB 252 CKECPKIHIDRIDKDDVEEOEHMRWLHEREIKDKMLIEFYESPDEPRRRRPFSGS 311
    |||||
OY 301 VNSKLSIKTLPMSLISGLTAGMLMTDAGRKLVTNTWYIGTLGCLMTWITKA 353
    |||||
```

Db 312 VNSKLSIKTKLPMSLILSGTAGMLMTDAGRKLYVNTWITYGTLGCLWYTIKA 364

```
|||||
RESULT 15
US-09-901-484A-5
; Sequence 5, Application US/09901484A
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate Cancer Gene
; FILE REFERENCE: GEN-T11XC3D2
; CURRENT APPLICATION NUMBER: US/09/901,484A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 08/996,306
; PRIOR FILING DATE: 1997-12-22
; PRIOR APPLICATION NUMBER: US 60/099,658
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: US 09/218,207
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: US 09/338,907
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: US 09/853,526
; PRIOR FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-901-484A-5
```

```
Query Match 100.0%; Score 353; DB 23; Length 364;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVVLGTAPRYVLANGVWRLLSAFLPARFYQALDDRLXYQOSMVLFFFEENTYG 60
DB 12 MRYLLPSVVLGTAPRYVLANGVWRLLSAFLPARFYQALDDRLXYQOSMVLFFFEENTYG 71
QY 61 VOILLYGDLPKNKENIYLANHOSYVDWIVADILAIQONALGHVRYVLEKGLKMLPLYGC 120
DB 72 VOILLYGDLPKNKENIYLANHOSYVDWIVADILAIQONALGHVRYVLEKGLKMLPLYGC 131
QY 121 YFAQHGIGIYVKSASFNEKEMRNKLQSYVDAGTPLYLIFPEGSTRYNPEQTKVLSASQAF 180
DB 132 YFAQHGIGIYVKSASFNEKEMRNKLQSYVDAGTPLYLIFPEGSTRYNPEQTKVLSASQAF 191
QY 181 AAQRGIAVLAKHVLTPRIKATHAFFDCMKNYLDAIYDVTVYVEGKDDGGQRRSPYTEPL 240
DB 192 AAQRGIAVLAKHVLTPRIKATHAFFDCMKNYLDAIYDVTVYVEGKDDGGQRRSPYTEPL 251
QY 241 CECEPKIHIIHIDRIKDVPEEOEHMRMLHERFEIKDKMLIEFYESPDERKRRPFGKS 300
DB 252 CECEPKIHIIHIDRIKDVPEEOEHMRMLHERFEIKDKMLIEFYESPDERKRRPFGKS 311
QY 301 VNSKLSIKTKLPMSLILSGTAGMLMTDAGRKLYVNTWITYGTLGCLWYTIKA 353
DB 312 VNSKLSIKTKLPMSLILSGTAGMLMTDAGRKLYVNTWITYGTLGCLWYTIKA 364
```

Search completed: August 28, 2002, 11:25:19  
Job time: 497 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:17:27 ; Search time 83.88 Seconds  
(without alignments)  
1018.645 Million cell updates/sec

Title: US-09-853-526-4

Perfect score: 353  
Sequence: 1 MRYLPSVVLGTAPTYVLA.....YVNTWIXTLLGLMTVTKA 353

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 747981 seqs, 242050750 residues

Word size : 0

Total number of hits satisfying chosen parameters: 747981

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Pending\_Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	353	100.0	364	US-10-184-648-63	Sequence 63, Appl
2	269	76.2	269	US-10-074-045-47	Sequence 47, Appl
3	252	71.4	353	US-09-629-469A-13028	Sequence 13028, A
4	9	2.5	135	US-10-184-648-52	Sequence 52, Appl
5	9	2.5	157	US-10-184-648-66	Sequence 66, Appl
6	9	2.5	240	US-09-895-913A-184	Sequence 184, Appl
7	8	2.3	236	US-09-053-853A-1538	Sequence 1538, App
8	2.3	262	5	US-09-791-537-5820	Sequence 5820, Ap
9	2.3	459	5	US-09-791-537-110883	Sequence 110883,
10	2.3	483	5	US-09-791-537-110642	Sequence 110642,
11	2.0	44	5	US-09-826-734A-142	Sequence 142, Appl
12	2.0	77	5	US-09-548-936C-17	Sequence 17, Appl
13	2.0	7	5	US-09-791-537-150979	Sequence 150979,
14	2.0	117	5	US-09-791-537-31675	Sequence 31675, A
15	2.0	126	5	US-09-791-537-63528	Sequence 63528, A
16	2.0	194	5	US-09-791-537-53250	Sequence 53250, A
17	2.0	195	6	US-10-164-966-9	Sequence 9, Appl
18	2.0	195	6	US-10-184-648-51	Sequence 51, Appl
19	2.0	195	6	US-10-184-648-67	Sequence 67, Appl
20	2.0	195	6	US-10-184-648-86	Sequence 86, Appl
21	2.0	208	1	PCT-US02-10824-132	Sequence 132, App
22	2.0	208	5	US-09-791-537-57367	Sequence 57367, A
23	2.0	208	5	US-09-791-537-99121	Sequence 99121, A
24	2.0	208	5	US-09-791-537-128069	Sequence 128069,
25	2.0	208	5	US-09-791-537-142901	Sequence 142901,
26	2.0	208	5	US-09-791-537-143774	Sequence 143774,

27	7	2.0	208	6	US-10-096-327-2	Sequence 2, Appl
28	7	2.0	208	6	US-10-179-131-7124	Sequence 7124, Ap
29	7	2.0	208	6	US-10-189-360-5	Sequence 5, Appl
30	7	2.0	208	6	US-10-189-360-7	Sequence 7, Appl
31	7	2.0	208	6	US-10-189-360-8	Sequence 8, Appl
32	7	2.0	208	6	US-10-189-360-18	Sequence 18, Appl
33	7	2.0	234	5	US-09-791-537-148811	Sequence 148811,
34	7	2.0	246	5	US-09-791-537-43519	Sequence 43519, A
35	7	2.0	250	5	US-09-791-537-67119	Sequence 67119, A
36	7	2.0	251	5	US-09-791-537-29660	Sequence 29660, A
37	7	2.0	253	6	US-10-153-881-22428	Sequence 22428, A
38	7	2.0	262	5	US-09-791-537-751	Sequence 751, App
39	7	2.0	263	6	US-10-155-851-18654	Sequence 18654, A
40	7	2.0	264	5	US-09-791-537-24942	Sequence 24942, A
41	7	2.0	269	6	US-10-053-853A-1244	Sequence 1244, Ap
42	7	2.0	270	6	US-10-211-364-1026	Sequence 1026, Ap
43	7	2.0	286	7	US-60-360-039-1733	Sequence 1733, Ap
44	7	2.0	311	7	US-60-360-039-22196	Sequence 22196, A
45	7	2.0	364	5	US-09-791-537-42339	Sequence 42339, A

#### ALIGNMENTS

RESULT 1  
US-10-184-648-63  
Sequence 63, Application US/10184648  
GENERAL INFORMATION:  
APPLICANT: Meyers, Rachel E.  
APPLICANT: Williamson, Mark  
APPLICANT: Tsai, Feng-Ying  
APPLICANT: Hunter, John J.  
APPLICANT: Macbeth, Kyle J.  
APPLICANT: Rudolph-Owen, Laura A.  
APPLICANT: Leiby, Kevin R.  
APPLICANT: Kapeller-Libermann, Rosana  
APPLICANT: Olandt, Peter J.  
TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE FAMILY MEMBERS AND USES THEREOF  
FILE REFERENCE: 10448-192001  
CURRENT APPLICATION NUMBER: US/10/184, 648  
CURRENT FILING DATE: 2002-06-27  
PRIOR APPLICATION NUMBER: US 09/815, 028  
PRIOR FILING DATE: 2001-03-22  
PRIOR APPLICATION NUMBER: PCT/US01/09358  
PRIOR FILING DATE: 2001-03-22  
PRIOR APPLICATION NUMBER: US 60/191, 964  
PRIOR FILING DATE: 2000-03-24  
PRIOR APPLICATION NUMBER: US 09/801, 220  
PRIOR FILING DATE: 2001-03-07  
PRIOR APPLICATION NUMBER: PCT/US01/07269  
PRIOR FILING DATE: 2001-03-07  
PRIOR APPLICATION NUMBER: US 60/187, 456  
PRIOR FILING DATE: 2000-03-07  
PRIOR APPLICATION NUMBER: US 09/816, 714  
PRIOR FILING DATE: 2001-03-23  
PRIOR APPLICATION NUMBER: PCT/US01/09468  
PRIOR FILING DATE: 2001-03-23  
PRIOR APPLICATION NUMBER: US 60/191, 865  
PRIOR FILING DATE: 2000-03-24  
PRIOR APPLICATION NUMBER: US 09/844, 948  
PRIOR FILING DATE: 2001-04-27  
PRIOR APPLICATION NUMBER: PCT/US01/13805  
PRIOR FILING DATE: 2001-04-27  
PRIOR APPLICATION NUMBER: US 60/200, 604  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: US 09/861, 164  
PRIOR FILING DATE: 2001-05-18  
PRIOR APPLICATION NUMBER: PCT/US01/16292  
PRIOR FILING DATE: 2001-05-18  
PRIOR APPLICATION NUMBER: US 60/205, 408  
PRIOR FILING DATE: 2000-05-19  
PRIOR APPLICATION NUMBER: US 09/883, 060  
PRIOR FILING DATE: 2001-06-15

```

; PRIOR APPLICATION NUMBER: PCT/US01/19138
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,079
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/962,678
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: PCT/US01/29963
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/235,044
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 09/973,457
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: US 60/238,849
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 10/072,285
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: PCT/US02/03736
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/267,494
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: US 09/817,910
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: PCT/US01/09633
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/192,092
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: US 09/842,528
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: PCT/US01/40607
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 60/199,500
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/882,836
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19543
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/211,730
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/882,872
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19153
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,077
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-184-648-63
```

Query Match 100.0%; Score 353; DB 6; Length 364;

Best Local Similarity 100.0%; Pred. No. 0; Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MRYLLPSVLLGTAPTYLVANGVRLSAPLPARYQALDRLCYVQSMVLFEEENTG 60
    |||||
DB 12 MRYLLPSVLLGTAPTYLVANGVRLSAPLPARYQALDRLCYVQSMVLFEEENTG 71
    |||||
QY 61 VOILLGDLPRKNENIYLANHSTVDWIVADILAIKRONALGHVRYVLEKGLMPLTGC 120
    |||||
DB 72 VOILLGDLPRKNENIYLANHSTVDWIVADILAIKRONALGHVRYVLEKGLMPLTGC 131
    |||||
QY 121 YFAHGGIYVRSKAFNEKEKRNKLOSVDAGTMYLVIPEEGTRYNEPQTKVLSASQAF 180
    |||||
DB 132 YFAHGGIYVRSKAFNEKEKRNKLOSVDAGTMYLVIPEEGTRYNEPQTKVLSASQAF 191
    |||||
QY 181 AAAGGLAVLKHLVTPRIKATHVAFDCKKNYLDATYDTVYVEGKGDDGQRESPTMTEFL 240
    |||||
DB 192 AAAGGLAVLKHLVTPRIKATHVAFDCKKNYLDATYDTVYVEGKGDDGQRESPTMTEFL 251
    |||||
```

```

QY 241 CKECPKHIIHIDRDKDVPEOEHRMRWLHEREIKDKMLIEFYESPDPERRRPFCKS 300
    |||||
DB 252 CKECPKHIIHIDRDKDVPEOEHRMRWLHEREIKDKMLIEFYESPDPERRRPFCKS 311
    |||||
QY 301 VNSKISTKTLPSMLISGLTAGMLMTDAGRKLYVNTYITGLGCLWVTIKA 353
    |||||
DB 312 VNSKISTKTLPSMLISGLTAGMLMTDAGRKLYVNTYITGLGCLWVTIKA 364
    |||||
```

RESULT 2

US-10-074-045-47

; Sequence 47, Application US/10074045

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PT221C1

; CURRENT APPLICATION NUMBER: US/10/074,045

; CURRENT FILING DATE: 2002-02-14

; Prior Application removed - See file wrapper or Palm

; NUMBER OF SEQ ID NOS: 74

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 47

; LENGTH: 269

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-074-045-47

Query Match 76.2%; Score 269; DB 6; Length 269;

Best Local Similarity 100.0%; Pred. No. 8,7e+268; Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 85 TVDMIVADILAIKRONALGHVRYVLEKGLMPLTGCYFAHGGIYVRSKAFNEKRNK 144
    |||||
DB 1 TVDMIVADILAIKRONALGHVRYVLEKGLMPLTGCYFAHGGIYVRSKAFNEKRNK 60
    |||||
QY 145 LQSYVDAGTMYLVIPEEGTRYNEPQTKVLSASQAFPAAGLAVLKHLVPRIKATHVAF 204
    |||||
DB 61 LQSYVDAGTMYLVIPEEGTRYNEPQTKVLSASQAFPAAGLAVLKHLVPRIKATHVAF 120
    |||||
QY 205 DCKKNYLDATYDTVYVEGKGDDGQRESPTMTEFLCKECPKHIIHIDRDKDVPEOE 264
    |||||
DB 121 DCKKNYLDATYDTVYVEGKGDDGQRESPTMTEFLCKECPKHIIHIDRDKDVPEOE 180
    |||||
QY 265 HMRWLHEREIKDKMLIEFYESPDPERRRPFCKSVNSKISTKTLPSMLISGLTAGM 324
    |||||
DB 181 HMRWLHEREIKDKMLIEFYESPDPERRRPFCKSVNSKISTKTLPSMLISGLTAGM 240
    |||||
QY 325 LMTDAGRKLYVNTYITGLGCLWVTIKA 353
    |||||
DB 241 LMTDAGRKLYVNTYITGLGCLWVTIKA 269
    |||||
```

RESULT 3

US-09-629-469A-13028

; Sequence 13028, Application US/09629469A

; GENERAL INFORMATION:

; APPLICANT: OTA, TOSHIO

; APPLICANT: ISOGAI, TAKAO

; APPLICANT: NISHIKAWA, TETSUO

; APPLICANT: HAYASHI, KOJI

; APPLICANT: SAITO, KAORU

; APPLICANT: YAMAMOTO, JUNICHI

; APPLICANT: ISHII, SHIZUKO

; APPLICANT: SUGIYAMA, TOMOYASU

; APPLICANT: WAKAMATSU, AI

; APPLICANT: NAGAI, KEIICHI

; APPLICANT: OTSUKI, TETSUJI

; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE

; FILE REFERENCE: 084335/0123

; CURRENT APPLICATION NUMBER: US/09/629,469A

; CURRENT FILING DATE: 2000-07-28

; PRIOR APPLICATION NUMBER: JP 1999-248036

```

; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 1999-300253
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183,322
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 19025
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13028
; LENGTH: 353
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-629-469A-13028
```

```

Query Match          71.4%; Score 252; DB 5; Length 353;
Best Local Similarity 99.7%; Pred. No. 3.3e-250;
Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```

QY 1 MRLLPSVVLGTAAPTYVLANGVWRLLSAFLPARFYQALDDRLCYVQSVLFEFENYTG 60
   |||||||
DB 1 MRLLPSVVLGTAAPTYVLANGVWRLLSAFLPARFYQALDDRLCYVQSVLFEFENYTG 60
   |||||||
QY 61 VQILLXGDLPRKNENIITVLANHSTVDWIVADILAIKRONALGHVRYLKGKMLPLYGC 120
   |||||||
DB 61 VQILLXGDLPRKNENIITVLANHSTVDWIVADILAIKRONALGHVRYLKGKMLPLYGC 120
   |||||||
QY 121 YFAOHGIIYKRSKAKNEKEMRNKLSYVDAGTPMYLVIFPEGTGRNPEQTKVLSAQAF 180
   |||||||
DB 121 YFAOHGIIYKRSKAKNEKEMRNKLSYVDAGTPMYLVIFPEGTGRNPEQTKVLSAQAF 180
   |||||||
QY 181 AARGLAIVLKHVLTPTKATHVAFDCMKNYLDAIVDTVYVEKDKDGGQRRSEPTMEFL 240
   |||||||
DB 181 AARGLAIVLKHVLTPTKATHVAFDCMKNYLDAIVDTVYVEKDKDGGQRRSEPTMEFL 240
   |||||||
QY 241 CKCPRKIHIDRIDKDDVEEOEHMRWMLHERFEIKDKMLIEFYESPDERRRKRPFGKS 300
   |||||||
DB 241 CKCPRKIHIDRIDKDDVEEOEHMRWMLHERFEIKDKMLIEFYESPDERRRKRPFGKS 300
   |||||||
QY 301 VNSKLSIKTLPMSMLISGLTFAGLMTDAGRKLIVNTWIYGTLLGLMTWITKA 353
   |||||||
DB 301 VNSKLSIKTLPMSMLISGLTFAGLMTDAGRKLIVNTWIYGTLLGLMTWITKA 353
   |||||||
```

```

RESULT 4
US-10-184-648-52
; Sequence 52, Application US/10184648
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Williamson, Mark
; APPLICANT: Tsai, Fong-Ying
; APPLICANT: Hunter, John J.
; APPLICANT: Macbeth, Kyle J.
; APPLICANT: Rudolph-Owen, Laura A.
; APPLICANT: Leiby, Kevin R.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Oladit, Peter J.
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE FAMILY MEMBERS AND USES THEREOF
; FILE REFERENCE: 10448-192001
; CURRENT FILING DATE: 2002-06-27
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: US 09/815,028
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: PCT/US01/09358
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: US 60/191,964
```

```

; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: US 09/801,220
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: PCT/US01/07269
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 60/187,456
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/816,714
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: PCT/US01/09468
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,865
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: US 09/844,948
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: PCT/US01/13805
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/200,604
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 09/861,164
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: PCT/US01/16292
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,408
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/883,060
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19138
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,079
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/962,678
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: PCT/US01/29963
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/235,044
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 09/973,457
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: US 60/238,849
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 10/072,285
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: PCT/US02/03736
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/267,494
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: US 09/817,910
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: PCT/US01/09633
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/192,092
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: US 09/842,528
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: PCT/US01/40607
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 60/199,500
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/882,836
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19543
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/211,730
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/882,872
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19153
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,077
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 52
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: consensus sequence
US-10-184-648-52

Query Match          2.5%; Score 9; DB 6; Length 135;
Best Local Similarity 100.0%; Pred. No. 0 67;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      157 LVIRPEGTR 165
        |||||
Db       85 LVIRPEGTR 93

RESULT 5
US-10-184-648-66
; Sequence 66, Application US/10184648
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Williamson, Mark
; APPLICANT: Tsai, Fong-Ying
; APPLICANT: Hunter, John J.
; APPLICANT: Macbeth, Kyle J.
; APPLICANT: Rudolph-Owen, Laura A.
; APPLICANT: Leiby, Kevin R.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Olendt, Peter J.
; TITLE OF INVENTION: NOVEL HUMAN TRANSERASE FAMILY MEMBERS AND USES THEREOF
; FILE REFERENCE: 10448-192001
; CURRENT APPLICATION NUMBER: US/10/184,648
; PRIOR FILING DATE: 2002-06-27
; PRIOR APPLICATION NUMBER: US 09/815,028
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: PCT/US01/09358
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: US 60/191,964
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: US 09/801,220
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: PCT/US01/07269
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 60/187,456
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/816,714
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: PCT/US01/09468
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,865
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: US 09/844,948
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: PCT/US01/13805
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/200,604
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 09/861,164
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: PCT/US01/16292
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,408
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/883,060
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19138
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,079
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/962,678
; PRIOR FILING DATE: 2001-09-25
```

```
; PRIOR APPLICATION NUMBER: PCT/US01/29963
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/235,044
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 09/973,457
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: US 60/238,849
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 10/072,285
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: PCT/US02/03736
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/267,494
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: US 09/817,910
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: PCT/US01/09653
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/192,092
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: US 09/842,528
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: PCT/US01/40607
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 60/199,500
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/882,836
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19543
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/211,730
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/882,872
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19153
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,077
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ IDS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 66
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: consensus sequence
US-10-184-648-66

Query Match          2.5%; Score 9; DB 6; Length 157;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      211 LDATYDVTY 219
        |||||
Db       4 LDATYDVTY 12

RESULT 6
US-09-895-913A-184
; Sequence 184, Application US/09895913A
; GENERAL INFORMATION:
; APPLICANT: Kleantous, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean Francois
; APPLICANT: Oomen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding Novel Helicobacter Polypeptides in the Helicobacter
; FILE REFERENCE: 06133/043002
; CURRENT APPLICATION NUMBER: US/09/895,913A
; PRIOR FILING DATE: 2001-06-29
```

;; PRIOR APPLICATION NUMBER: US 08/881,227  
;; PRIOR FILING DATE: 1997-06-24  
;; NUMBER OF SEQ ID NOS: 368  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 184  
;; LENGTH: 240  
;; TYPE: PRT  
;; ORGANISM: Helicobacter pylori  
US-09-895-913A-184

Query Match 2.5%; Score 9; DB 5; Length 240;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 LVIFPEGTR 165  
DB 142 LVIFPEGTR 150

RESULT 7  
US-10-053-853A-1538  
;; Sequence 1538, Application US/10053853A  
;; GENERAL INFORMATION:  
;; APPLICANT: HAYASHI, Hideo  
;; APPLICANT: SHINAGAWA, Hideo  
;; APPLICANT: MAKINO, Kozo  
;; APPLICANT: HAYASHI, Tetsuya  
;; APPLICANT: OHNISHI, Makoto  
;; APPLICANT: HATTORI, Masahira  
;; APPLICANT: KUROKAWA, Ken

;; TITLE OF INVENTION: Polynucleotide molecules and polypeptides specific to Enterohemo  
;; FILE REFERENCE: 2002-00604/MMC/01704  
;; CURRENT APPLICATION NUMBER: US/10/053, 853A  
;; CURRENT FILING DATE: 2002-05-28  
;; PRIOR APPLICATION NUMBER: JP2001-112010  
;; PRIOR FILING DATE: 2001-01-24  
;; NUMBER OF SEQ ID NOS: 1866  
;; SEQ ID NO 1538  
;; LENGTH: 236  
;; TYPE: PRT  
;; ORGANISM: Escherichia coli O157:H7  
US-10-053-853A-1538

Query Match 2.3%; Score 8; DB 6; Length 236;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 LTPRIKAT 200  
DB 227 LTPRIKAT 234

RESULT 8  
US-09-791-537-5820  
;; Sequence 5820, Application US/09791537  
;; GENERAL INFORMATION:  
;; APPLICANT: Bionomix, Inc.  
;; APPLICANT: Debe, Derek  
;; APPLICANT: Danzer, Joseph  
;; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB  
;; TITLE OF INVENTION: METHODS OF USE THEREOF  
;; FILE REFERENCE: 261/210  
;; CURRENT APPLICATION NUMBER: US/09/791,537  
;; CURRENT FILING DATE: 2001-02-22  
;; NUMBER OF SEQ ID NOS: 153055  
;; SOFTWARE: PatentIn version 3.0  
;; SEQ ID NO 5820  
;; LENGTH: 262  
;; TYPE: PRT  
;; ORGANISM: Chlamydia trachomatis

US-09-791-537-5820

Query Match 2.3%; Score 8; DB 5; Length 262;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 HVLPRIK 198  
DB 202 HVLPRIK 209

RESULT 9  
US-09-791-537-110883  
;; Sequence 110883, Application US/09791537  
;; GENERAL INFORMATION:  
;; APPLICANT: Bionomix, Inc.  
;; APPLICANT: Debe, Derek  
;; APPLICANT: Danzer, Joseph  
;; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME  
;; TITLE OF INVENTION: METHODS OF USE THEREOF  
;; FILE REFERENCE: 261/210  
;; CURRENT APPLICATION NUMBER: US/09/791,537  
;; CURRENT FILING DATE: 2001-02-22  
;; NUMBER OF SEQ ID NOS: 153055  
;; SOFTWARE: PatentIn version 3.0  
;; SEQ ID NO 110883  
;; LENGTH: 459  
;; TYPE: PRT  
;; ORGANISM: Bacillus subtilis  
US-09-791-537-110883

Query Match 2.3%; Score 8; DB 5; Length 459;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 167 NPEQTKVL 174  
DB 292 NPEQTKVL 299

RESULT 10  
US-09-791-537-110642  
;; Sequence 110642, Application US/09791537  
;; GENERAL INFORMATION:  
;; APPLICANT: Bionomix, Inc.  
;; APPLICANT: Debe, Derek  
;; APPLICANT: Danzer, Joseph  
;; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME  
;; TITLE OF INVENTION: METHODS OF USE THEREOF  
;; FILE REFERENCE: 261/210  
;; CURRENT APPLICATION NUMBER: US/09/791,537  
;; CURRENT FILING DATE: 2001-02-22  
;; NUMBER OF SEQ ID NOS: 153055  
;; SOFTWARE: PatentIn version 3.0  
;; SEQ ID NO 110642  
;; LENGTH: 483  
;; TYPE: PRT  
;; ORGANISM: Pisum sativum  
US-09-791-537-110642

Query Match 2.3%; Score 8; DB 5; Length 483;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 AGTPMIVL 158  
DB 184 AGTPMIVL 191

RESULT 11

```
US-09-826-734A-142
; Sequence 142, Application US/09826734A
; GENERAL INFORMATION:
; APPLICANT: Fernandes, Elma R
; APPLICANT: Mishra, Vishnu S
; APPLICANT: Leach, Martin D
; APPLICANT: Shinkets, Richard A
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Kekuda, Kamesh
; TITLE OF INVENTION: Novel Polynucleotides And Polypeptides Encoded Thereby
; FILE REFERENCE: 15966-754
; CURRENT APPLICATION NUMBER: US/09/826,734A
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/195,576
; PRIOR FILING DATE: 2000-04-26
; NUMBER OF SEQ ID NOS: 264
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 142
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-826-734A-142
```

```
Query Match          2.0%; Score 7; DB 5; Length 44;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 159 IFEPECTR 165
      |||||||
      3 IFEPECTR 9
```

```
RESULT 12
; Sequence 17, Application US/09548936C
; GENERAL INFORMATION:
; APPLICANT: Gopez, Leonel Jorge
; APPLICANT: Saras, Jan
; APPLICANT: Claesson-Welsh, Lena
; APPLICANT: Heidln, Carl-Henrik
; TITLE OF INVENTION: PTPLI BINDING AGENTS
; FILE REFERENCE: L0461/7084
; CURRENT APPLICATION NUMBER: US/09/548,936C
; CURRENT FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: US 09/100,804
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: US 08/596,291
; PRIOR FILING DATE: 1994-09-01
; PRIOR APPLICATION NUMBER: US 08/115,573
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patentln Version 3.0
; SEQ ID NO 17
; LENGTH: 77
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-548-936C-17
```

```
Query Match          2.0%; Score 7; DB 5; Length 77;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 125 HGGIYVK 131
      |||||||
      20 HGGIYVK 26
```

```
RESULT 13
US-09-791-537-150979
; Sequence 150979, Application US/09791537
; GENERAL INFORMATION:
```

```
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patentln version 3.0
; SEQ ID NO 150979
; LENGTH: 96
; TYPE: PRT
; ORGANISM: pdb 3PDZA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: X is an unknown amino acid
US-09-791-537-150979
```

```
Query Match          2.0%; Score 7; DB 5; Length 96;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 125 HGGIYVK 131
      |||||||
      32 HGGIYVK 38
```

```
RESULT 14
US-09-791-537-31675
; Sequence 31675, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patentln version 3.0
; SEQ ID NO 31675
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-791-537-31675
```

```
Query Match          2.0%; Score 7; DB 5; Length 117;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 125 HGGIYVK 131
      |||||||
      43 HGGIYVK 49
```

```
RESULT 15
US-09-791-537-63528
; Sequence 63528, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
```

; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 63528  
; LENGTH: 126  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-791-537-63528

Query Match 2.0%; Score 7; DB 5; Length 126;  
Best Local Similarity 100.0%; Pred. No. 71;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 125 HGGIYVK 131  
| | | | |  
Db 52 HGGIYVK 58

Search completed: August 28, 2002, 11:26:50  
Job time: 563 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:15:47 ; Search time 39.16 Seconds  
(without alignments)  
866.178 Million cell updates/sec

Title: US-09-853-526-4

Perfect score: 353

Sequence: 1 MRYLPSVVLGTAPRYVLA.....YVNTWIXFTLLGLMTYTKA 353

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 283138 seqs, 96089334 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	2.5	237	2	B71827
2	9	2.5	240	2	D64688
3	9	2.5	264	2	E82903
4	9	2.5	918	2	T34057
5	8	2.3	83	1	W8BPG7
6	8	2.3	83	1	W8BPG7
7	8	2.3	236	2	H90892
8	8	2.3	236	2	B87524
9	8	2.3	239	2	A85725
10	8	2.3	262	2	A71478
11	8	2.3	282	2	G71334
12	8	2.3	396	2	SS4641
13	8	2.3	411	2	C87586
14	8	2.3	424	2	T49607
15	8	2.3	458	2	G83735
16	8	2.3	459	2	E63793
17	8	2.3	483	2	T06459
18	8	2.3	611	2	A85000
19	8	2.0	54	2	S68122
20	7	2.0	57	2	D86108
21	7	2.0	117	2	I81209
22	7	2.0	126	2	I81210
23	7	2.0	141	2	SS6117
24	7	2.0	143	2	B90523
25	7	2.0	148	2	C86732
26	7	2.0	194	2	PC1136
27	7	2.0	195	2	C83854
28	7	2.0	208	1	A38432
29	7	2.0	208	1	A41914

30	7	2.0	208	1	JC1409	heparin-binding Eg
31	7	2.0	209	2	B83329	probable acyltrans
32	7	2.0	210	2	A10342	probable acyltrans
33	7	2.0	211	2	E70476	2-acylglycerophos
34	7	2.0	225	2	H95244	conserved hypothet
35	7	2.0	226	2	E98109	conserved hypothet
36	7	2.0	230	2	E81397	probable 1-acylgly
37	7	2.0	240	2	S75162	hypothetical prote
38	7	2.0	241	2	B97019	1-acyl-sn-glycerol
39	7	2.0	246	2	T46546	hypothetical prote
40	7	2.0	247	2	A81957	1-acylglycerol-3-p
41	7	2.0	247	2	G81013	1-acyl-sn-glycerol
42	7	2.0	250	1	A31757	homeotic protein H
43	7	2.0	250	2	T27772	26S proteasome reg
44	7	2.0	250	2	E70104	1-acylglycerol-3-p
45	7	2.0	250	2	AF3384	outer membrane pro

#### ALIGNMENTS

RESULT 1  
B71827  
probable 1-acylglycerol-3-phosphate O-acyltransferase (EC 2.3.1.51) - Helicobacter pylori  
C:Species: Helicobacter pylori  
A:Variate: strain J99  
C>Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 06-Oct-2000  
C:Accession: B71827  
R:Alm, R.A.; Ling, L.S.L.; Molr, D.T.; King, B.L.; Brown, E.D.; Doly, P.C.; Smith, D.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jlang, O.; Taylor, D.E.; Vovis, G.F. Nature 397, 176-180, 1999  
A>Title: Genomic sequence comparison of two unrelated isolates of the human gastric p  
A:Reference number: A71800; MID:99120557  
A:Accession: B71827  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-237 <ARN>  
A:Cross-references: GB:AE001550; GB:AE001439; MID:g4155672; PIDN:AAD06852.1; PID:g415  
A:Experimental source: strain J99  
C:Genetics:  
A:Gene: p1sc  
C:Superfamily: mouse 1-acylglycerol-3-phosphate O-acyltransferase  
C:Keywords: acyltransferase; coenzyme A

Query Match 2.5%; Score 9; DB 2; Length 237;  
Best Local Similarity 100.0%; Pred. No. 0.39;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 LVTFPEGTR 165  
DB 142 LVTFPEGTR 150

RESULT 2  
D64688  
probable 1-acylglycerol-3-phosphate O-acyltransferase (EC 2.3.1.51) - Helicobacter pylori  
C:Species: Helicobacter pylori  
C>Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 06-Oct-2000  
C:Accession: D64688  
R:Tomb, J.F.; White, O.; Kariavase, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R. Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKee son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watney, Nature 388, 539-547, 1997  
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, A.; Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A:Reference number: A64520; MID:97394467  
A:Accession: D64688  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-240 <TOM>  
A:Cross-references: GB:AE000636; GB:AE000511; MID:g2314517; PIDN:AAD08393.1; PID:g231  
C:Superfamily: mouse 1-acylglycerol-3-phosphate O-acyltransferase

C:keywords: acyltransferase; coenzyme A

Query Match 2.5%; Score 9; DB 2; Length 240;  
Best Local Similarity 100.0%; Pred. No. 0.39;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 157 LVIFPECTR 165  
DB 142 LVIFPECTR 150

RESULT 3  
E82903  
1-acyl-sn-glycerol-3-phosphate acyltransferase U0344 [Imported] - Ureaplasma urealyticum  
C:Species: Ureaplasma urealyticum  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C:Accession: E82903  
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Helner, C.R.; Chen, E.Y.; Cassell, G.H.  
submitted to GenBank, February 2000  
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min  
A:Reference number: A82870  
A:Accession: E82903  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-264 <GLA>  
A:Cross-references: GB:AE002131; GB:AF222894; NID:g6899316; PIDN:AAF30753.1; GSPDB:GN001  
A:Experimental source: serovar 3; biovar 1  
C:Genetics:  
A:Gene: pISC: U0344  
A:Genetic code: SGC3

Query Match 2.5%; Score 9; DB 2; Length 264;  
Best Local Similarity 100.0%; Pred. No. 0.43;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 157 LVIFPECTR 165  
DB 156 LVIFPECTR 164

RESULT 4  
T34057  
hypothetical protein F28B3.5 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
C:Accession: T34057  
R:Geisel, C.; Kramer, J.; Smith, A.  
submitted to the EMBL Data Library, May 1997  
A:Description: The sequence of C. elegans cosmid F28B3.  
A:Reference number: Z21469  
A:Accession: T34057  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-918 <CEI>  
A:Cross-references: EMBL:AF003136; PIDN:AB93636.1; GSPDB:GN00019; CESP:F28B3.5  
A:Experimental source: strain Bristol N2; clone F28B3  
C:Genetics:  
A:Gene: CESP:F28B3.5  
A:Map position: 1  
A:introns: 85/3; 129/3; 235/3; 418/2; 482/3; 532/3; 736/3; 829/3; 850/2

Query Match 2.5%; Score 9; DB 2; Length 918;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 211 LDAIYDVT 219  
DB 762 LDAIYDVT 770

RESULT 5  
W8BPG7  
gene 18.7 protein - phage T7  
C:Species: phage T7  
C:Date: 13-Jun-1983 #sequence\_revision 13-Jun-1983 #text\_change 23-Jul-1999  
C:Accession: A04413; S42337  
R:Dunn, J.U.; Thompson, K.  
submitted to the Nucleic Acid Sequence Database, September 1982  
A:Reference number: A94615  
A:Accession: A04413  
A:Molecule type: DNA  
A:Residues: 1-83 <DUN>  
R:Dunn, J.U.; Studier, F.W.  
J. Mol. Biol. 166, 477-535, 1983  
A:Title: Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7  
A:Reference number: S42283; MUID:83241725  
A:Accession: S42337  
A:Molecule type: DNA  
A:Residues: 1-83 <DUN>  
A:Cross-references: EMBL:V01146; NID:g431187; PIDN:CAA24439.1; PID:g15615  
A:Note: the authors did not translate the codon for residue 1  
C:Genetics:  
A:Gene: 18.7  
A:Map position: 92.73-93.35  
C:Superfamily: phage T7 gene 18.7 protein

Query Match 2.3%; Score 8; DB 1; Length 83;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 305 LSIKKTLP 312  
DB 22 LSIKKTLP 29

RESULT 6  
W8BPT3  
gene 18.7 protein - phage T3  
C:Species: phage T3  
A:Note: host Escherichia coli  
C:Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 23-Jul-1999  
C:Accession: E23476  
R:Yamada, M.; Fujisawa, H.; Kato, H.; Hamada, K.; Minagawa, T.  
Virology 151, 350-361, 1986  
A:Title: Cloning and sequencing of the genetic right end of bacteriophage T3 DNA.  
A:Reference number: A94339; MUID:86209997  
A:Accession: E23476  
A:Molecule type: DNA  
A:Residues: 1-83 <YAM>  
A:Cross-references: GB:M14784; NID:g215810; PIDN:AA92527.1; PID:g1196765  
R:Yamada, M.; Fujisawa, H.; Kato, H.; Hamada, K.; Minagawa, T.  
Virology 154, 246, 1986  
A:Reference number: A94344  
A:Contents: annotation; erratum; corrections to coding regions  
C:Genetics:  
A:Gene: 18.7  
C:Superfamily: phage T7 gene 18.7 protein

Query Match 2.3%; Score 8; DB 1; Length 83;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 305 LSIKKTLP 312  
DB 22 LSIKKTLP 29

RESULT 7  
H90892  
probable fimbrial chaperone protein precursor [Imported] - Escherichia coli (strain O  
C:Species: Escherichia coli

C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001  
C:Accession: H90892  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Iishi, K.; Yokoyama, K.; Han, C.G.  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and gene  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: H90892  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-236 <HAY>  
A:Cross-references: GB:BA000007, PIDN:BA835535.1; PID:g13361578; GSPDB:GN00154  
A:Experimental source: strain O157:H7, substrain RMD 0509952  
C:Genetics:  
A:Gene: ECS2112  
C:Superfamily: chaperone protein papp

Query Match 2.3%; Score 8; DB 2; Length 236;  
Best Local Similarity 100.0%; Pred. No. 4.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 193 LFPRIKAT 200  
|||||  
Db 227 LFPRIKAT 234

RESULT 8  
B87524  
1-acyl-sn-glycerol-3-phosphate acyltransferase [Imported] - *Caulobacter crescentus*  
C:Species: *Caulobacter crescentus*  
C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
C:Accession: B87524  
R:Nierman, W.C.; Feldlyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
B.; Land, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwin, M.L.; Haft, D.H.; Kolon  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete genome sequence of *Caulobacter crescentus*.  
A:Reference number: A87249; MUID:21173698; PMID:11259647  
A:Accession: B87524  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-236 <STO>  
A:Cross-references: GB:AE005673; NID:g13423724; PIDN:AAK24190.1; GSPDB:GN00148  
C:Genetics:  
A:Gene: CC2219

Query Match 2.3%; Score 8; DB 2; Length 236;  
Best Local Similarity 100.0%; Pred. No. 4.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 158 VIFPECTR 165  
|||||  
Db 146 VIFPECTR 153

RESULT 9  
A85725  
Probable fimbrial chaperone protein Z2201 [Imported] - *Escherichia coli* (strain O157:H7,  
C:Species: *Escherichia coli*  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
C:Accession: A85725  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
Hiller, L.; Grothbeck, E.J.; Davis, N.W.; Lm, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: A85725  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-239 <STO>  
A:Cross-references: GB:AE005174; NID:g12515164; PIDN:AAG56261.1; GSPDB:GN00145; UMG:222

A:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: Z2201  
C:Superfamily: chaperone protein papp

Query Match 2.3%; Score 8; DB 2; Length 239;  
Best Local Similarity 100.0%; Pred. No. 4.3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 193 LFPRIKAT 200  
|||||  
Db 230 LFPRIKAT 237

RESULT 10  
A71478  
probable metal dependent hydrolase - *Chlamydia trachomatis* (serotype D, strain UW3/Cx  
C:Species: *Chlamydia trachomatis*  
C:Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 08-Oct-1999  
C:Accession: A71478  
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitche  
Science 282, 754-759, 1998  
A:Title: Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia t*  
A:Reference number: A71570; MUID:99000809  
A:Accession: A71478  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-262 <ARN>  
A:Cross-references: GB:AE001344; GB:AE001273; NID:g3329188; PIDN:AAC68333.1; PID:g332  
A:Experimental source: serotype D, strain UW-3/Cx  
C:Genetics:  
A:Gene: YycJ

Query Match 2.3%; Score 8; DB 2; Length 262;  
Best Local Similarity 100.0%; Pred. No. 4.6;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 191 HVLPRIK 198  
|||||  
Db 202 HVLPRIK 209

RESULT 11  
G71334  
probable lysophosphatidic acid acyltransferase - *Syphilis spirochete*  
C:Species: *Treponema pallidum* subsp. *pallidum* (*Syphilis spirochete*)  
C:Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 05-Nov-1999  
C:Accession: G71334  
R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; G  
rison, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; M  
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.  
Science 281, 375-388, 1998  
A:Title: Complete genome sequence of *Treponema pallidum*, the *syphilis spirochete*.  
A:Reference number: A71250; MUID:98332770  
A:Accession: G71334  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-282 <COL>  
A:Cross-references: GB:AE001215; GB:AE000520; NID:g3322631; PIDN:AAC65346.1; PID:g332  
A:Experimental source: strain Nichols  
C:Genetics:  
A:Gene: TP0361

Query Match 2.3%; Score 8; DB 2; Length 282;  
Best Local Similarity 100.0%; Pred. No. 5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 158 VIFPECTR 165  
|||||  
Db 170 VIFPECTR 177

## RESULT 12

S54641

probable membrane protein YDR018c - yeast (*Saccharomyces cerevisiae*)N:Alternate names: hypothetical protein D3246; hypothetical protein P2F396; hypothetical C:Species: *Saccharomyces cerevisiae*

C&gt;Date: 08-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 20-Jun-2000

C:Accession: S54641; S63425; S67831; S72116

R:Edman, K.; Brown, D.; Hamlyn, N.; Bowman, S.

submitted to the EMBL Data Library, May 1995

A:Reference number: S54638

A:Accession: S54641

A:Molecule type: DNA

A:Residues: 1-396 &lt;DED&gt;

A:Cross-references: EMBL:249770; NID:g840867; PIDN:CAA9843.1; PID:g840871

A:Experimental source: strain AB972

R:Elde, L.G.; Sander, C.; Prydz, H.

submitted to the EMBL Data Library, February 1996

A:Description: Sequencing and analysis of a 35.4 kb region on the left arm of chromosome

A:Reference number: S63416

A:Accession: S63425

A:Molecule type: DNA

A:Residues: 1-396 &lt;EID&gt;

A:Cross-references: EMBL:X95966; NID:g1216215; PIDN:CAA65210.1; PID:g1216225

R:Prydz, H.; Elde, L.G.

submitted to the Protein Sequence Database, July 1996

A:Reference number: S67822

A:Accession: S67831

A:Molecule type: DNA

A:Residues: 1-396 &lt;PRY&gt;

A:Cross-references: EMBL:274314; NID:g1431443; PIDN:CAA9838.1; PID:g1431444; MIPS:YDR01

A:Experimental source: strain S288C

R:Elde, L.G.; Sander, C.; Prydz, H.

Yeast 12, 1085-1090, 1996

A:Title: Sequencing and analysis of a 35.4 kb region on the left arm of chromosome IV fr

A:Reference number: S72107; MIMD:97051598

A:Accession: S72116

A:Molecule type: DNA

A:Status: nucleic acid sequence not shown; translation not shown

A:Residues: 1-396 &lt;EIM&gt;

A:Cross-references: EMBL:X95966; NID:g1216215; PIDN:CAA65210.1; PID:g1216225

A:Note: The nucleotide sequence was submitted to the EMBL Data Library, February 1996

C:Genetics:

A:Map position: 4R

A:Note: YDR018c

C:Superfamily: probable membrane protein YBR042c

C:Keywords: transmembrane protein

F:27-43/Domain: transmembrane #status predicted &lt;TM1&gt;

F:59-85/Domain: transmembrane #status predicted &lt;TM2&gt;

F:376-392/Domain: transmembrane #status predicted &lt;TM3&gt;

Query Match

Best Local Similarity 100.0%; Pred. No. 6.8;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 211 LDAIYDVT 218

DB 265 LDAIYDVT 272

## RESULT 13

C87586

metal ion efflux membrane fusion protein family [imported] - *Caulobacter crescentus*C:Species: *Caulobacter crescentus*

C&gt;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 30-Jun-2001

C:Accession: C87586

R:Merman, W.C.; Feldblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gilm, M.L.; Haft, D.H.; Kolton

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of *Caulobacter crescentus*.

A:Reference number: A87249; MIMD:21173698; PMID:11259647

A:Accession: C87586

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-411 &lt;STO&gt;

A:Cross-references: GB:AE005673; NID:g13424311; PIDN:AAK24667.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC2722

C:Superfamily: nickel-cobalt resistance determinant structural protein CnrB; 11poy1/b

Query Match

Best Local Similarity 100.0%; Pred. No. 7;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 316 ILSGTAG 323

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

DB 380 ILSGTAG 387

QY 167 NPEQTKVL 174  
| | | | |  
Db 291 NPEQTKVL 298

Search completed: August 28, 2002, 11:20:08  
Job time: 261 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:09:02 ; Search time 75.33 Seconds  
(without alignments)  
520.497 Million cell updates/sec

Title: US-09-853-526-4  
Perfect score: 353  
Sequence: 1 MRYLPSVLLGTAPRYVLA.....YVNTWYTGILLGLMTWTKA 353

Scoring table:  
Gapop 60.0 , Gapext 60.0

Searched: 747574 seqs, 111073796 residues

Word size : 0

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

A.Geneseq\_032802:\*  
1: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:\*  
2: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:\*  
3: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:\*  
4: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:\*  
5: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:\*  
6: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:\*  
7: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:\*  
8: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:\*  
9: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:\*  
10: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT:\*  
11: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:\*  
12: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:\*  
13: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:\*  
14: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:\*  
15: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:\*  
16: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:\*  
17: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:\*  
18: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:\*  
19: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:\*  
20: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:\*  
21: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:\*  
22: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	353	100.0	353	22	AAB85532
2	353	100.0	353	22	AAB85552
3	353	100.0	364	20	AAV36729
4	353	100.0	364	22	AAV50128
5	353	100.0	364	22	AAU00665
6	352	99.7	352	20	AAV36728
7	290	82.2	353	22	AAV36591
8	269	76.2	269	22	AAE09653
9	252	71.4	353	22	AAV36595
10	232	71.4	372	22	AAV41377
11	230	65.2	291	20	AAV36744

12	228	64.6	228	20	AAV36740
13	199	56.4	261	20	AAV36745
14	185	52.4	185	20	AAV36753
15	185	52.4	315	20	AAV36751
16	184	52.1	300	20	AAV36752
17	154	43.6	182	20	AAV36750
18	116	32.9	238	20	AAV36743
19	85	24.1	96	20	AAV36749
20	70	19.8	257	22	ABG01360
21	70	19.8	980	22	ABG11977
22	62	17.6	66	20	AAV36748
23	62	17.6	68	20	AAV36747
24	62	17.6	77	20	AAV36742
25	53	17.6	90	20	AAV36746
26	53	15.0	97	22	ABG12868
27	53	15.0	1104	22	ABG12107
28	32	9.1	354	20	AAV36741
29	9	2.5	9	20	AAV36734
30	9	2.5	157	22	AAV50130
31	9	2.5	237	18	AAV55540
32	9	2.5	240	19	AAV98477
33	9	2.5	242	18	AAV53441
34	8	2.3	9	20	AAV36737
35	8	2.3	239	22	ABB52710
36	8	2.3	396	21	AAV99484
37	8	2.3	557	22	ABB71470
38	8	2.3	1820	22	ABB62370
39	7	2.0	12	16	AAV87705
40	7	2.0	12	20	AAV43149
41	7	2.0	65	16	AAV87720
42	7	2.0	82	16	AAV43164
43	7	2.0	82	16	AAV87718
44	7	2.0	82	16	AAV87719
45	7	2.0	82	20	AAV43160

#### ALIGNMENTS

RESULT 1	
AAB85532	standard; protein; 353 AA.
XX	
AC	AAB85532;
XX	
DT	25-SEP-2001 (first entry)
XX	
DE	Human secreted protein (clone id HLTH084) .
XX	
XX	
KW	Secreted protein; immunosuppressive; antiarthritic; antirheumatic;
KW	antiproliferative; cytostatic; cardiant; vasotrophic; cerebroprotective;
KW	neotropic; neuroprotective; antibacterial; virucide; fungicide; human;
KW	optimallogical; gene therapy.
XX	
OS	Homo sapiens.
XX	
PN	WO200155430-A1.
XX	
PD	02-AUG-2001.
XX	
PF	17-JAN-2001; 2001MO-US01431.
XX	
PR	31-JAN-2000; 2000US-0179065.
PR	04-FEB-2000; 2000US-0180628.
PR	12-SEP-2000; 2000US-0231968.
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR, Olsen HS;
PI	Moore PA, Wei P, Edner R, Duan DR, Shi Y, Choi GH, Fiscella M;
PI	Ni J, Ruben SM, Barash SC;
XX	
DR	WPI; 2001-476220/51.

Truncated PGI prot  
PGI splice variant  
PGI splice variant  
PGI splice variant  
PGI splice variant  
PGI splice variant  
PGI splice variant  
PGI splice variant  
Novel human diagno  
Novel human diagno  
Novel human diagno  
PGI splice variant  
PGI splice variant  
PGI splice variant  
Novel human diagno  
Novel human diagno  
Novel human diagno  
Mouse PGI protein  
PGI box 3 motif.  
Lyso phosphatidic a  
H. pylori ORF 12ge  
H. pylori GHPO 903  
H. pylori ORF 06gp  
Acyl glycerol tran  
Escherichia coli p  
Yeast acyltransfer  
Drosophila melanog  
Coconut AG3PAT try  
Coconut 1-acylglyc  
Coconut LPAAT clon  
1-acylglycerol-3-p  
Coconut LPAAT clon  
Coconut LPAAT clon  
1-acylglycerol-3-p

DR N-PSDB: AAH46942.  
XX  
PT 17 isolated nucleic acid molecules encoding human secreted proteins,  
XX used to preventing, treating or ameliorating a medical condition -  
XX  
PS Claim 11: Page 450-451: 482pp: English.  
XX  
XX The invention provides novel human secreted proteins and polynucleotides  
CC encoding them. The secreted proteins can be expressed by standard  
CC recombinant methodology. The secreted proteins and polynucleotides are  
CC used to prevent, treat or ameliorate a medical condition in e.g. humans,  
CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They can  
CC also be used in diagnosing a pathological condition. The antibodies to  
CC the proteins can also be used in alleviating symptoms associated with the  
CC disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme  
CC linked immunoassays (ELISA). Disorders which are diagnosed or  
CC treated include autoimmune diseases e.g. rheumatoid arthritis,  
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,  
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
CC e.g. cerebral ischemia, angiodenesis, nervous system disorders e.g.  
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and  
CC ocular disorders e.g. corneal infection. The polypeptides can also be  
CC used to aid wound healing and epithelial cell proliferation, to prevent  
CC skin aging due to sunburn, to maintain organs before transplantation, for  
CC supporting cell culture of primary tissues, to regenerate tissues and in  
CC chemotaxis. The polypeptides can also be used as a food additive or  
CC preservative to increase or decrease storage capabilities. The present  
CC sequence represents a human secreted protein.  
XX  
SQ Sequence 353 AA:  
  
Query Match 100.0%; Score 353; DB 22; Length 353;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MRYLLPSVNLGTAPTVYVLMGWRLSAFLPARYOALDRLCYQSWLFEFFENTG 60  
DB 1 mryllpsvnlgtaptvylawgwrrllsalfparfygaldrlcyqswmlfffenytg 60  
QY 61 VOILLYDLPKRNENIYLANHOSTVDMIVADILAIRONALGHVRYVLEKGLMWPLXGC 120  
DB 61 vqlllygdlpknenillylanhgstvdwivadilairqnalghvryvlekgkwlplxgc 120  
QY 121 YFAOHGSIYVKRSKAKFNEKEMRNKLOSYVDAGTPMYLVIREGTRNPEQTKVLSASQAF 180  
DB 121 yfaohgsllyvkrsakfnekemrnlqsyvdagtpmylviregtrnpeqtkvlsasqaf 180  
QY 181 AAOGLAVLKHVLTPTRIKATHVAFDCKKNYLDIAYDTVYVEGDDGQRRSEPTMTFEL 240  
DB 181 aagglavllkhvltprtkathvafdcmknyldaiydtvvyegkddgqrrseptmtefl 240  
QY 241 CKECPKTHIHIDRIKDDVPEOEHRMRMLHERFEIKDKMLIEFYESPDPERRKRFPKGS 300  
DB 241 ckecpkthihdridkddvpeegehrmrmlherfelkdkmliefesppderkrfpkgs 300  
QY 301 VNSKLSIKTKLPMSMLISGLTAGMLMTDAGRKLTVNTWIVGTLLGCLMWYTIKA 353  
DB 301 vnskslisktklpmsmlisgltagmlmtdagrkltvntwivgtllgclmwtyika 353  
  
RESULT 2  
AAB85352  
ID AAB85552 standard; protein: 353 AA.  
XX  
AC AAB85552:  
XX  
XX 25-SEP-2001 (first entry)  
XX  
XX Human secreted protein (clone ID HSLIA81).  
XX  
XX Secreted protein: immunosuppressive; antiarthritic; antirheumatic;  
KM antiproliferative; cytostatic; cardiac; vasotropic; cerebroprotective;

KW neurotropic; neuroprotective; antibacterial; virucide; fungicide; human;  
KM ophthalmological; gene therapy.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO20015430-A1.  
PN  
XX  
XX 02-AUG-2001.  
PD  
XX  
XX 17-JAN-2001; 2001MO-US01431.  
PF  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 12-SEP-2000; 2000US-0231968.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA  
XX  
XX Rosen CA, Komatsu G, Baker KP, Birse CE, Soppet DR, Olsen HS;  
PI Moore PA, Wei P, Eber R, Duan DR, Shi Y, Choi GH, Fiscella M;  
PI Ni J, Ruben SM, Barash SC;  
XX  
XX WPI: 2001-476220/51.  
DR N-PSDB: AAH46962.  
XX  
XX  
XX 17 isolated nucleic acid molecules encoding human secreted proteins,  
PT used to preventing, treating or ameliorating a medical condition -  
XX  
XX  
PS Claim 11: Page 465-466; 482pp: English.  
XX  
XX  
XX The invention provides novel human secreted proteins and polynucleotides  
CC encoding them. The secreted proteins can be expressed by standard  
CC recombinant methodology. The secreted proteins and polynucleotides are  
CC used to prevent, treat or ameliorate a medical condition in e.g. humans,  
CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They can  
CC also be used in diagnosing a pathological condition. The antibodies to  
CC the proteins can also be used in alleviating symptoms associated with the  
CC disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme  
CC linked immunoassays (ELISA). Disorders which are diagnosed or  
CC treated include autoimmune diseases e.g. Rheumatoid arthritis,  
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,  
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
CC e.g. cerebral ischemia, angiodenesis, nervous system disorders e.g.  
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and  
CC ocular disorders e.g. corneal infection. The polypeptides can also be  
CC used to aid wound healing and epithelial cell proliferation, to prevent  
CC skin aging due to sunburn, to maintain organs before transplantation, for  
CC supporting cell culture of primary tissues, to regenerate tissues and in  
CC chemotaxis. The polypeptides can also be used as a food additive or  
CC preservative to increase or decrease storage capabilities. The present  
CC sequence represents a human secreted protein.  
XX  
SQ Sequence 353 AA:  
  
Query Match 100.0%; Score 353; DB 22; Length 353;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MRYLLPSVNLGTAPTVYVLMGWRLSAFLPARYOALDRLCYQSWLFEFFENTG 60  
DB 1 mryllpsvnlgtaptvylawgwrrllsalfparfygaldrlcyqswmlfffenytg 60  
QY 61 VOILLYDLPKRNENIYLANHOSTVDMIVADILAIRONALGHVRYVLEKGLMWPLXGC 120  
DB 61 vqlllygdlpknenillylanhgstvdwivadilairqnalghvryvlekgkwlplxgc 120  
QY 121 YFAOHGSIYVKRSKAKFNEKEMRNKLOSYVDAGTPMYLVIREGTRNPEQTKVLSASQAF 180  
DB 121 yfaohgsllyvkrsakfnekemrnlqsyvdagtpmylviregtrnpeqtkvlsasqaf 180  
QY 181 AAOGLAVLKHVLTPTRIKATHVAFDCKKNYLDIAYDTVYVEGDDGQRRSEPTMTFEL 240  
DB 181 aagglavllkhvltprtkathvafdcmknyldaiydtvvyegkddgqrrseptmtefl 240



QY	241	CKECKHIIHIDRIDKRVPEEGEEMRWLHERFIEKIKMKLIEFESDPERKKRPKGS	300
DB	241	CKECPKIIHIDRIDKRVPEEGEEMRWLHERFIEKIKMKLIEFESDPERKKRPKGS	300
QY	301	VNSKSIKTKLPMSLITSGITAGMIMTGTAGRLYNTWYTGILGCLWYTIKA	353
DB	301	VNSKSIKTKLPMSLITSGITAGMIMTGTAGRLYNTWYTGILGCLWYTIKA	353
RESULT	3		
ID	AAV36729	standard; Protein: 364 AA.	
XX	AAV36729		
AC	AAV36729;		
XX			
DT	27-SEP-1999	(first entry)	
XX			
DE	Human PGI protein sequence.		
XX			
KW	PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;		
KW	cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.		
XX			
OS	Homo sapiens.		
XX			
PN	W09932644-A2.		
PD			
XX	01-JUL-1999.		
XX			
PF	22-DEC-1998;	98WO-IB02133.	
XX			
PR	09-SEP-1998;	98US-0099658.	
PR	22-DEC-1997;	97US-0996306.	
XX			
PA	(GEST ) GENSET.		
XX			
PI	Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;		
XX			
DR	WPI: 1999-405178/34.		
DR	N-PSDB: AAZ00870.		
XX			
PT	Use of a prostate cancer associated gene and biallelic markers		
XX	derived from it		
XX			
PS	Claim 7; Page 190-191; 385pp; English.		
XX			
CC	The invention relates to a mammalian PGI gene and protein, and a set of		
CC	PGI biallelic markers. The PGI polynucleotide and biallelic markers are		
CC	used in a hybridisation assay, a sequencing assay, or in an		
CC	allele-specific amplification assay for determining the identity of a		
CC	nucleotide at a PGI-related biallelic marker. The methods can be used to		
CC	detect and to assess the risk of developing cancer or prostate cancer.		
CC	Early-stage diagnosis of prostate cancer relies on prostate specific		
CC	antigen (PSA) dosage. However, the effectiveness of this is limited due		
CC	to its inability to discriminate between malignant and non-malignant		
CC	affections of the organ. A need exists for both a reliable diagnostic		
CC	procedure which would enable early-stage diagnosis, and for preventative		
CC	and curative treatments of the disease. The PGI gene can be used for		
CC	detection of prostate cancer, and the risk of developing it in the		
CC	future, and can also be used to determine therapies for the disease.		
XX			
XX			
SO	Sequence 364 AA;		
QY	Query Match	100.0%; Score 353; DB 20; Length 364;	
DB	Best Local Similarity	100.0%; Pred. No. 0;	
Matches	353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1	MRYLPSVILGTAFYVILANGVWRLTASPLPARYQALDRLXCVYQSMVFFPENTG	60
DB	12	mrYlpsvIlgtApCyIaWgVwRlTsPlPaRyQaLdRlXcvYqSmVlFfEnTg	71
QY	61	VOILLYGLDPNNKRNIIIVLANHGSTVDWIVDDIILRONALGHVRYVLEKGLKWPILYGC	120

Db	72	vqlllygdipknkenillylaangstcvwiyadilalrqnalghvrylkqgikwlpipygc	131
Oy	121	YFAOHGGIYVRSRKSFKNEKEKRNKLQSVYAGTPMVLVIPEEGTRVNEOTKVLASQAF	180
Db	132	yfaqbgqylgvyrnsakfknekmrnklqsvyadqfpmylvlfpqgtrlyrpeqtkvlaaaqaf	191
Oy	181	AAQGLAVLKHVLPRIKATHVAFDCKNKNYLDALVDYTVVYEGKDDGGQRRESPTMTFEL	240
Db	192	aagrglavlkhvlprlkathvafdcmknyldalvdvtvvyeqkddggqrresptmtefl	251
Oy	241	CKECKTHIHDRIDKKDVPEQEHMRRLHERETIDKMLIERYESPDPERKKRPFGKS	300
Db	252	ckecpkthihdriddkdvpeeqehmrtrwherfelxdkmliefyespdperkrfpqks	311
Oy	301	VNSKLSIKTLPMSLISGLFAGMLAMDAGKTLVNWVIYGTLLGLCMTVRIKA	353
Db	312	vnsklsiktlpmsllisglfagmlmdagrklyvntviygtllglcmtvrika	364
RESULT 4			
AA50128	ID	AA50128 standard; Protein; 364 AA.	
XX	AC	AA50128;	
XX	DT	21-DEC-2001 (first entry)	
DE	XX	Human acyltransferase 27417.	
XX	XX		
KW	KW	Acyltransferase 27417; human; cell proliferation; cell migration;	
KW	KW	cell differentiation; ovary cancer; brain cancer; colon cancer;	
KW	KW	lung cancer; tumour; metastasis; sarcoma; carcinoma;	
KW	KW	adenocarcinoma; antitumour; lipid metabolism; diagnosis; therapy.	
XX	XX		
OS	OS	Homo sapiens.	
XX	XX		
FT	FT	Key	Location/Qualifiers
FT	FT	Peptide	1..49
FT	FT	Protein	/label= Signal_peptide
FT	FT	Domain	/label= Mature_protein
FT	FT	Domain	/label= N-terminal_domain
FT	FT	Domain	/label= Transmembrane_domain
FT	FT	Domain	/label= C-terminal_domain
FT	FT	Domain	/label= Acyltransferase_domain
FT	FT	Modified-site	/note= "asn is N-glycosylated"
FT	FT	Modified-site	/note= "O-phosphorylated by cAMP- and cGMP-dependent protein kinase"
FT	FT	Modified-site	11..13
FT	FT	Modified-site	/note= "O-phosphorylated by protein kinase C"
FT	FT	Modified-site	144..146
FT	FT	Modified-site	/note= "O-phosphorylated by protein kinase C"
FT	FT	Modified-site	205..207
FT	FT	Modified-site	/note= "O-phosphorylated by protein kinase C"
FT	FT	Modified-site	317..319
FT	FT	Modified-site	/note= "O-phosphorylated by protein kinase C"
FT	FT	Modified-site	361..363
FT	FT	Modified-site	/note= "O-phosphorylated by protein kinase C"
FT	FT	Modified-site	95..98
FT	FT	Modified-site	/note= "O-phosphorylated by casein kinase II"
FT	FT	Modified-site	158..161
FT	FT	Modified-site	/note= "O-phosphorylated by casein kinase II"
FT	FT	Modified-site	246..249
FT	FT	Modified-site	/note= "O-phosphorylated by casein kinase II"
FT	FT	Modified-site	82..89
FT	FT	Modified-site	/note= "O-phosphorylated by tyrosine kinase"

FT	Modified-site	219..226	/note="O-phosphorylated by tyrosine kinase"
FT	Modified-site	23..28	/note="N-myristoylated"
FT	Modified-site	130..135	/note="N-myristoylated"
FT	Modified-site	330..335	/note="N-myristoylated"
FT	Modified-site	352..357	/note="N-myristoylated"
FT	Modified-site	230..343	/note="Amidation"
FT	Peptide	105..126	/note="predicted leucine zipper"
PN	MO200173051-A2.		
PD	04-OCT-2001.		
XX	26-MAR-2001; 2001WO-US09633.		
PR	24-MAR-2000; 2000US-192092P.		
XX	(MILL-) MILLENNIUM PHARM INC.		
PA	Meyers RA, Rudolph-Owen L, Macbeth KJ, Tsai F, Williamson M;		
PI	WPI; 2001-626267/72.		
DR	N-PSDB; AAH27094.		
XX			
FT	New polynucleotide encodes a polypeptide for the diagnosis and		
FT	treatment of cancer and to alter cellular lipid metabolism comprises a		
FT	polynucleotide encoding human acyltransferases designated 46743 and		
PT	27417		
XX			
PS	Claim 9; Fig 8A-B; 150pp; English.		
XX			
CC	The present sequence is that of a novel human protein, termed		
CC	27417, which shows the structural characteristics of members of the		
CC	acyltransferase family, including an acyltransferase domain. The		
CC	invention provides novel acyltransferase 27417 nucleic acids and		
CC	polypeptides, as well as methods for detecting their presence, and		
CC	methods for screening for compounds that modulate their expression		
CC	or activity. Such compounds can be used to treat conditions		
CC	related to aberrant activity or expression of 27417 protein or		
CC	nucleic acid, such as conditions involving deficient cellular		
CC	proliferation, migration and/or differentiation. Also provided are		
CC	methods for inhibiting the proliferation or migration, or inducing		
CC	the killing, of a 27417-expressing cell, e.g. a hyperproliferative		
CC	and/or metastatic cell, by contact with a compound that modulates		
CC	the activity or expression of the 27417 protein or nucleic acid.		
CC	The 27417-expressing cell is found in the healthy or diseased		
CC	heart, blood vessels, kidney, skeletal muscle, brain or liver, or		
CC	especially in a solid tumour, a soft tissue tumour or a metastatic		
CC	lesion, a sarcoma, a carcinoma or an adenocarcinoma, and in		
CC	particular in a hyperproliferative and/or metastatic cell found in		
CC	ovarian, brain, colon or lung cancer. The compounds may also be		
CC	used to modulate lipid metabolism in a 27417-expressing cell.		
CC	Methods of disease diagnosis, e.g. by determining the		
CC	presence of a genetic alteration in a 27417 polypeptide, and for		
CC	evaluating the efficacy of a treatment of a disorder, are also		
CC	provided.		
XX			
SO	Sequence 364 AA;		
0Y	Query Match	100.0%; Score 353; DB 22; Length 364;	
	Best Local Similarity	100.0%; Pred. No. 0;	
db	Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
	1 MRYLLPSVLLGTAFTYVLAWGVMKLSAFIPARYQALDRLYCYVQSMVLFPEENYTG 60		
	12 mryllpsvllgtaftcyvlawgvmlklsafiparfygaldtdlcyvqsgmvlffennytg 71		

QY	61	VQILLEGDLPENKNNKNTIYYLANHSTVDWIVADLIARONALGHVRYVLEKGLKMLPIYGC	120
QY	72	VQILLEGDLPENKNNKNTIYYLANHSTVDWIVADLIARONALGHVRYVLEKGLKMLPIYGC	131
QY	121	YFAOHGGTYVRSKAKFNPEKEMRNKLGSTVYAGAPMWLVIFPEEGRXNPEOTVLSASQAF	180
Db	132	YFAOHGGTYVRSKAKFNPEKEMRNKLGSTVYAGAPMWLVIFPEEGRXNPEOTVLSASQAF	191
QY	181	AAORGLAVLKHVLPRIKATHVAEDCMKNLYDAIYDVTVVYEGKDDGGORRESPTMTEFL	240
Db	192	AAQGLAVLKHVLPRIKATHVAEDCMKNLYDAIYDVTVVYEGKDDGGORRESPTMTEFL	251
QY	241	CKECPKTHIHTRDRKDDVPEEGEHNRRMLHEEFELKDKMLIEFYSPSPPERKRRPPGKS	300
Db	252	CKECPKTHIHTRDRKDDVPEEGEHNRRMLHEEFELKDKMLIEFYSPSPPERKRRPPGKS	311
QY	301	VNSKLSIKTLPSPMLISGLTAGMLMTDAGRKLYVTWTVYGTLLGLCAWTYIKA	353
Db	312	VNSKLSIKTLPSPMLISGLTAGMLMTDAGRKLYVTWTVYGTLLGLCAWTYIKA	364
RESULT	5		
AAU00665			
ID	AAU00665	standard; Protein; 364 AA.	
AC	AAU00665;		
XX			
DT	07-SEP-2001	(first entry)	
XX			
DE	Human lysophosphatidic acid acyltransferase isoform LPAAT-epsilon.		
XX			
KM	Lysophosphatidic acid acyltransferase epsilon; LPAAT-epsilon; LPA; PA;		
KM	lysophosphatidic acid; phosphatidic acid; acylation; cellular activation;		
KW	phospholipid signalling; mitogenesis; inflammation; autoimmune disease;		
KW	oncology; cancer; obesity; gene therapy.		
XX			
OS	Homo sapiens.		
XX			
PN	MO200134782-A1.		
PD	17-MAY-2001.		
XX			
PF	02-NOV-2000; 2000MO-US30193.		
XX			
PR	09-NOV-1999; 99US-0436919.		
XX			
PA	(CELL-) CELL THERAPEUTICS INC.		
XX			
P1	Leung DW;		
XX			
DR	WPI: 2001-335920/35.		
DR	N-PSDB: AAS00649.		
XX			
PT	Novel isolated human isoform of lysophosphatidic acid		
PT	acyltransferase-epsilon useful for diagnostic, therapeutic and		
XX	screening purposes		
XX			
PS	Claim 1; Fig 1; 48pp; English.		
XX			
CC	The sequence represents a human lysophosphatidic acid acyltransferase		
CC	(LPAAT) isoform, LPAAT-epsilon. LPAAT catalyses the acylation of		
CC	lysophosphatidic acid (LPA) to phosphatidic acid (PA). LPA and PA have		
CC	been identified as phospholipid signalling molecules that affect a wide		
CC	range of biological responses. PA is involved in cellular activation and		
CC	mitogenesis. Compounds that block PA generation and hence diminish lipid		
CC	biosynthesis and the signal involved in cell activation are of		
CC	therapeutic interest in the areas of inflammation and oncology (e.g.		
CC	autoimmune diseases and cancer) as well as obesity treatment.		
CC	LPAAT-epsilon and its corresponding DNA can be used in screening assays		
CC	to detect agents that stimulate or inhibit the activity of LPAAT and,		
CC	therefore, PA. The DNA is useful in tests to detect the presence or		
CC	expression of LPAAT-epsilon in relation to certain diseases and		

CC conditions, and in disease prevention and treatment. The sequences of the  
CC invention are also useful for diagnosis of diseases and conditions in  
CC which the expression of LPAAT enzyme is abnormal.

XX Sequence 364 AA;

Query Match 100.0%; Score 353; DB 22; Length 364;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRYLLPSVLLGTAPTYVLAMGWRLLSAFLPAFYQALDRLCYQSWLFEFFENYTG 60  
DB 12 mryllpsvllgtaptyvlamgwrrllsalfparfyqalddrlcyqswmlfffenytg 71  
OY 61 VOILLYGDLPKRNENITTYLANHOSYVDWIVADILAIRONALGHVRYLKEGLKWLPLYGC 120  
DB 72 vqilllygdlpknenillylanhgstvdwivadilaatrgnalghvrylkeglkwlplygc 131  
OY 121 YFAOHGIYVKRSKAFKNEKEMRNKIOSYVDAGTPMYLVIPPEGTRNYPEQTKVLSAQAF 180  
DB 132 yfaohgiyvkrsakfnekemrnlkgsyvdagtpmylvippegtrnypeqtkvlsasqaf 191  
OY 181 AAOGRGLAVLKHVLTLPRIKATHVAFDCMKNYLDALYDVTVVYEGKDGQGRRESPTWTEFL 240  
DB 192 aagrglavlkhvltprikathvafdcmknyldalylvdvtyvegkdqgrresptwtefl 251  
OY 241 CKCEPKIHIIHIDRIDKDVDEOEHRMRWLHERPEIKDKMLIEFYESPDERRRKRRFPKGS 300  
DB 252 ckcepkihihidridkdvdeoehrmrwlherpeikdkmliefyespderrrkrrfpkgs 311  
OY 301 VNSKLSIKKTLPMSLISGLTAGMLMTDAGRKLTVNNTWIVYGTLLGCLMTVITKA 353  
DB 312 vnsksikktlpmsllsgltagmlmtdagrklyvntwivgtllgclwtvika 364

## RESULT 6

AA36728  
ID AAY36728 standard; Protein; 352 AA.

AC AAY36728;

DT 27-SEP-1999 (first entry)

DE Human PGI protein sequence.

KW PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;  
KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.

XX Homo sapiens.

XX WO932644-A2.

XX 01-JUL-1999.

XX 22-DEC-1998; 98WO-1B02133.

XX 09-SEP-1998; 98US-0099658.

XX 22-DEC-1997; 97US-0996306.

XX (GENSET ) GENSET.

PI Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;

DR WPI; 1999-405178/34.

DR N-PSDB; AA200872.

XX Use of a prostate cancer associated gene and biallelic markers  
PT derived from it

XX Claim 7; Page 190; 385pp; English.

XX The invention relates to a mammalian PGI gene and protein, and a set of

CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
CC used in a hybridisation assay, a sequencing assay, or in an  
CC allele-specific amplification assay for determining the identity of a  
CC nucleotide at a PGI-related biallelic marker. The methods can be used to  
CC detect and to assess the risk of developing cancer or prostate cancer.  
CC Early-stage diagnosis of prostate cancer relies on prostate specific  
CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
CC to its inability to discriminate between malignant and non-malignant  
CC affections of the organ. A need exists for both a reliable diagnostic  
CC procedure which would enable early-stage diagnosis, and for preventative  
CC and curative treatments of the disease. The PGI gene can be used for  
CC detection of prostate cancer, and the risk of developing it in the  
CC future, and can also be used to determine therapies for the disease.

XX Sequence 352 AA;

Query Match 99.7%; Score 352; DB 20; Length 352;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 352; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRYLLPSVLLGTAPTYVLAMGWRLLSAFLPAFYQALDRLCYQSWLFEFFENYTG 60  
DB 1 mryllpsvllgtaptyvlamgwrrllsalfparfyqalddrlcyqswmlfffenytg 60  
OY 61 VOILLYGDLPKRNENITTYLANHOSYVDWIVADILAIRONALGHVRYLKEGLKWLPLYGC 120  
DB 61 vqilllygdlpknenillylanhgstvdwivadilaatrgnalghvrylkeglkwlplygc 120  
OY 121 YFAOHGIYVKRSKAFKNEKEMRNKIOSYVDAGTPMYLVIPPEGTRNYPEQTKVLSAQAF 180  
DB 121 yfaohgiyvkrsakfnekemrnlkgsyvdagtpmylvippegtrnypeqtkvlsasqaf 180  
OY 181 AAOGRGLAVLKHVLTLPRIKATHVAFDCMKNYLDALYDVTVVYEGKDGQGRRESPTWTEFL 240  
DB 181 aagrglavlkhvltprikathvafdcmknyldalylvdvtyvegkdqgrresptwtefl 240  
OY 241 CKCEPKIHIIHIDRIDKDVDEOEHRMRWLHERPEIKDKMLIEFYESPDERRRKRRFPKGS 300  
DB 241 ckcepkihihidridkdvdeoehrmrwlherpeikdkmliefyespderrrkrrfpkgs 300  
OY 301 VNSKLSIKKTLPMSLISGLTAGMLMTDAGRKLTVNNTWIVYGTLLGCLMTVITKA 352  
DB 301 vnsksikktlpmsllsgltagmlmtdagrklyvntwivgtllgclwtvika 352

## RESULT 7

AAM39591  
ID AAM39591 standard; Protein; 353 AA.

AC AAM39591;

DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 2736.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinesis; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.

XX Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US34263.

XX 21-JAN-2000; 2000US-0488725.

XX 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.  
PR 19-JUL-2000; 2000US-0620312.  
PR 03-AUG-2000; 2000US-0653450.  
PR 14-SEP-2000; 2000US-0662191.  
PR 19-OCT-2000; 2000US-0693036.  
PR 29-NOV-2000; 2000US-0727344.  
XX (HYSE-) HYSEQ INC.  
XX  
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX  
XX WPI: 2001-442253/47.  
DR N-PSDB: AA158747.  
XX  
XX Novel nucleic acids and polypeptides, useful for treating disorders  
PT such as central nervous system injuries -  
XX  
XX Example 4: SEQ ID NO 2736; 10078bp; English.  
XX  
XX The invention relates to human nucleic acids (AA157798-AA161369) and  
CC the encoded polypeptides (AA158642-AA162213) with nootropic.  
CC Immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.  
XX  
XX Sequence 353 AA:  
SQ

Query Match 82.2%; Score 290; DB 22; Length 353;  
Best Local Similarity 100.0%; Pred. No. 3.8e-287;  
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTAFTYVLAWGVRLSAFLPARFYQALDRLCYVQSMYLPFFENTYG 60  
Db 1 mryllpsvlllgtaftylvawgvrlsafflparfyqalddrlcyvqsmlylffenytg 60  
QY 61 VOILLYDLPKKNENIYYLANHSTVDWIVADILAIKONALGHVRYLKEGLKMLPLYGC 120  
Db 61 vqilllygdldpknkeniyylanhstvdwivadilairqnalghvrylkeglkwlplygc 120  
QY 121 YFAHGIGIYVRSKFNKEKEMRNKLOSYVDAGTPMYLYIPEEGRYNDEQKVVYASQAF 180  
Db 121 yfahgigiyvrsakfnkekemrnklosyvdaqtpmylylpeegcrynpeqtkvysaqaf 180  
QY 181 AAGRGVLVLRKAVLPRIKATVAFDCMKNYLDATYDTVYVEGKDDGQRRSESTMTIEFL 240  
Db 181 aagrgvlvlrkavlprikatvafdcmknyldatydvtvyvegkddgqrrsestlmtefl 240  
QY 241 CKECPRKHIDRIKDKVPEDEQHMRRLHEFEIKDKMLIEFYESPDP 290  
Db 241 ckecpkhlhldridkdkvpeeqhmrmlhefeikdkmliefyespdp 290

RESULT 8  
AAE09653  
ID AAE09653 standard; Protein: 269 AA.  
XX  
AC AAE09653;  
XX  
DT 19-NOV-2001 (first entry)

XX  
DE Human gene 15 encoded lipid metabolism protein HHA5B32, SEQ ID NO:47.  
XX  
XX  
KW Human; lipid metabolism protein; nootropic; neuroprotective; cardiatic;  
KW cerebroprotective; hepatotropic; antidiabetic; ophthalmic; nephrotropic;  
KW immune disorder; autoimmune disease; rheumatoid arthritis; glossitis;  
KW systemic lupus erythematosus; conjunctivitis; inflammatory disorder;  
KW respiratory disorder; asthma; allergy; CNS disorder; Alzheimer's disease;  
KW Parkinson's disease; atherosclerosis; cardiovascular disorder; cancer;  
KW coronary disease; familial hypercholesterolaemia; hyperlipidaemia;  
KW hemapoleptic disorder; hypolipidaemia; lipidosi; Gaucher's disease;  
KW Tay-sach's disease; mental retardation; gene therapy; antisense therapy.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO20015203-A1.  
XX  
XX 02-AUG-2001.  
PD  
XX  
XX 17-JAN-2001; 2001WO-US01327.  
PF  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218230.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226688.  
PR 22-AUG-2000; 2000US-0227182.  
PR 22-AUG-2000; 2000US-0227009.  
PR 23-AUG-2000; 2000US-0228924.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 06-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.

PR 14-SEP-2000; 2000US-0233397.  
PR 14-SEP-2000; 2000US-0233398.  
PR 14-SEP-2000; 2000US-0233399.  
PR 14-SEP-2000; 2000US-0233400.  
PR 14-SEP-2000; 2000US-0233401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.

PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM.  
XX  
XX WPI; 2001-45716/49.  
XX N-PSDB; AAD16620.  
XX  
XX  
XX Isolated lipid metabolism polypeptide for screening to identify  
XX antagonists and agonists that may enhance or block activities mediated  
XX by lipid metabolism proteins and also for testing and detection e.g.  
XX diagnosis -  
XX  
XX  
XX Claim 11; SEQ ID No 47; 453pp; English.  
XX  
XX AAD16606-AAD16627 represent cDNAs corresponding to human lipid metabolism  
XX protein genes, and AAD09639-AAD09660 represent the proteins they encode.  
XX AAD16628-AAD16647 represent human lipid metabolism genomic DNAs. The  
XX genes and their corresponding proteins are useful for preventing,  
XX treating or ameliorating medical conditions such as immune disorders,  
XX inflammatory disorders (e.g., conjunctivitis, glossitis), autoimmune  
XX diseases (e.g., rheumatoid arthritis, systemic lupus erythematosus),  
XX respiratory disorders (asthma, allergy), CNS disorders (e.g., Alzheimer's  
XX disease, Parkinson's disease), haematopoietic disorders, cardiovascular  
XX disorders (e.g., atherosclerosis, coronary disease), hyperlipidaemias,  
XX hypolipidaemia, lipidoses, familial hypercholesterolaemia, Gaucher's  
XX disease, Tay-sachs's disease, diabetes mellitus, cancer and other enzyme  
XX abnormalities. They are used for the prevention of mental retardation  
XX and the enlargement of spleen and liver. They are used in gene therapy  
XX and antisense therapy. The present sequence represents a human lipid  
XX metabolism protein of the invention.  
XX CC Note: The sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX  
XX Sequence 269 AA:  
SQ  
Query Match 76.2%; Score 269; DB 22; Length 269;  
Best Local Similarity 100.0%; Pred No. 7.7e-266;  
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 85 TVDMIVADILAIRONALGHVRYVLKGLKWLPLGYCYFAQHGSIYVRSARKEKEMRNK 144  
DB 1 tvdmivadilairgnalghvryvlkeglkwlplgygyfaqhgslkykrskkfnekmrnk 60  
QY 145 LOSYVAGTPMYLVITPEEGRIYRPEOTKYLASAQAFAAGGLAVLKHVLTPRTKATHAVR 204  
DB 61 lgsyvdaqtpmylvitpfeqtrlynpetkylsasqafaagglavlvkhvltprtkathavf 120  
QY 205 DCMKNYLDATYDVTVYEGGDGGRRESPTMEPLCKECPKIHIIHIDRDKDVPDEOE 264  
DB 121 dcmknyldatydvtvyeggdggrresptmeflckecpkihihidridkdvpeee 180  
QY 265 HMRRLWHERFEIKDKMLIEFYESPDPERRRKRPFGKSVNSKLSIKTKLPSPMLISGLTAGM 324  
DB 181 hmrrwlherfeikdkmliefyespdperrrkrfpgkavnskslkslklpsmlisgltagm 240  
QY 325 LMTDAGRKLYVNTWITGLTGLWVITKA 353  
DB 241 lmtdagrklyvntwiltglglwvltika 269

RESULT 9  
 AAB93595  
 ID AAB93595 standard; Protein; 353 AA.  
 AC AAB93595;  
 XX  
 XX AAB93595;  
 XX  
 XX 26-JUN-2001 (first entry)  
 DT  
 DE Human protein sequence SEQ ID NO:13028.  
 DE  
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.  
 XX  
 OS Homo sapiens.  
 PN  
 XX EPI074617-A2.  
 PD  
 XX 07-FEB-2001.  
 XX  
 XX 28-JUL-2000; 2000EP-0116126.  
 XX  
 XX 29-JUL-1999; 99JP-0248036.  
 PR 27-AUG-1999; 99JP-0300253.  
 PR 11-JAN-2000; 2000JP-018776.  
 PR 02-MAY-2000; 2000JP-0183767.  
 PR 09-JUN-2000; 2000JP-0241899.  
 XX  
 PA (HELI-) HELIX RES INST.  
 XX  
 XX Ota T, Isogai T, Nishikawa T, Hayashi K, Salto K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 DR WPI: 2001-318749/34.  
 XX  
 XX Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs -  
 XX  
 PS Claim 8; SEQ ID 13028; 2537pp + CD ROM; English.  
 XX  
 XX The present invention describes primer sets for synthesizing 5602  
 CC full-length cDNAs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and  
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.  
 XX  
 XX Sequence 353 AA;  
 XX

```

Db      1 mryllpsvlllgtcptyvllawgvrllsalfiparfylgalddrtlycyvgsamwlffienytg 60
Oy      61 voilllgedlpennkeniiylanhnostydmivadiilairqnalghvrvylkeglkmlplvyc 120
Db      61 vqlllygdlprkeniiaiylnhgsctvdwlvadilaalrqnalghvrvylkeglkmlplvyc 120
Oy      121 yFAOHGCIYVVRSAKFFNEKEKRNKLGOSYVAGCPMVLVIFPESTRKRNPEQTVLSASQAF 180
Db      121 yfaqgaglyvkrtsakfnekemrkvsgyvdaqfpmylvlifpegrtynpeqtvlasaqaf 180
Oy      181 AAORCLAWLKVHLVLPRIKATHTVAFDCKMKNVLDAIYQVTVVYVEGKDGDGGGRRESPTMTEPL 240
Db      181 aagqglavlkvhlvlprikachvafdcmknyldaiydvrvvyegkdgdgggrresptmtefl 240
Oy      241 CKECPKIHIDRIDKDDVEDEQHMRMRWLHERFEIKDKMLIEFYESPDPERRRRFPFGKS 300
Db      241 ckecpkhbihidriddkddveeqgmhrtrwlherfelkdkmllefyespdperrrrfpqks 300
Oy      301 VNSKISIKKTLPSMLISGLTAGMLMTDAGRKLYVNTWTVIYIGTLGCLWVTYTA 353
Db      301 vnsksikktlpsmlisgltagmlmtdagrklvntwtviygtllgclwtvltika 353

RESULT_10
AAM41377
ID      AAM41377 standard; Protein; 372 AA.
AC
AC      AAM41377;
CC
CC      22-OCT-2001 (first entry)
DE
DE      Human polypeptide SEQ ID NO 6308.
XX
XX
XX      Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW      peripheral nervous system; neuropathy; central nervous system; CNS;
KW      Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW      amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW      chemokine; thrombolytic; drug screening; arthritis; inflammation;
KW      leukaemia.
XX
XX
XX      Homo sapiens.
OS
OS      WO200153312-A1.
PN
PN      26-JUL-2001.
PD
PD      26-JUL-2001.
XX
XX
XX      26-DEC-2000; 2000MO-US34263.
PF
PF      21-JAN-2000; 2000US-048725.
XX
XX      25-APR-2000; 2000US-0532317.
PR      09-JUL-2000; 2000US-0598042.
PR      19-JUL-2000; 2000US-0620312.
PR      03-AUG-2000; 2000US-0653450.
PR      14-SEP-2000; 2000US-0662191.
PR      19-OCT-2000; 2000US-0693036.
PR      29-NOV-2000; 2000US-0727344.
XX
XX
XX      (HYSE-) HYSEQ INC.
PA
PA
XX      Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI      Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI      Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX      WPI: 2001-442253/47.
DR      N-PSDB; AAI60533.
XX
XX      Novel nucleic acids and polypeptides, useful for treating disorders
XX      such as central nervous system injuries -
XX
XX      Example 2; SEQ ID NO 6308; 10078bp; English.
XX
XX      The invention relates to human nucleic acids (AA157798-AA161369) and
CC      the encoded polypeptides (AAM38642-AAM42213) with noctropic,

```

CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC activation/inhibition activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukemias and  
 CC C.N.S disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.

XX Sequence 372 AA;

Query Match 71.4%; Score 252; DB 22; Length 372;

Best Local Similarity 99.7%; Pred. No. 2.3e-248;

Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRLILPSVVLGTRPTVYVLAAGWVRLLSARLPARFYALDDRLCYVQSWVLEFFENYTG 60  
 |||||||  
 DB 20 MYLILPSVVLIGTAPVYVLAAGWVRLLSARLPARFYALDDRLCYVQSWVLEFFENYTG 79  
 |||||||  
 QY 61 VOILLYGDLPRKNENIYYLANHOSTVDIVADILAIRONALGHVRYLKEGILKWLPLYGC 120  
 |||||||  
 DB 80 VQILLYGDLPKHENIYYLANHSTVDIVADILAIRONALGHVRYLKEGILKWLPLYGW 139  
 |||||||  
 QY 121 YFAOHGIVYKRSKAFNEKEKMRNKLQSYVDAGTPMYLVIEPEGRYRPEQKRVLSASQAF 180  
 |||||||  
 DB 140 YFAGHGIVYKRSKAFNEKEKMRNKLQSYVDAGTPMYLVIEPEGRYRPEQKRVLSASQAF 199  
 |||||||  
 QY 181 AAGRGVLAHVLTLPRIKATHVAFDCMKNYDAIYDVTVVYEGKDDGGRRESPTMTFEL 240  
 |||||||  
 DB 200 AAGRGVLAHVLTLPRIKATHVAFDCMKNYDAIYDVTVVYEGKDDGGRRESPTMTFEL 259  
 |||||||  
 QY 241 CKECPKIHIDRIDKDVPEEGEHRMRWLHERFEIKDKMLIEFESDPERRRRFPFGKS 300  
 |||||||  
 DB 260 CKECPKIHIDRIDKDVPEEGEHRMRWLHERFEIKDKMLIEFESDPERRRRFPFGKS 319  
 |||||||  
 QY 301 VNSKLSTIKTLPMSLILSGLTAGMLMTDAGRKLYVNTWIVYGTLLGCLMTWITKA 353  
 |||||||  
 DB 320 VNSKLSTIKTLPMSLILSGLTAGMLMTDAGRKLYVNTWIVYGTLLGCLMTWITKA 372

RESULT 11

AAV36744

ID AAV36744 standard; Protein; 291 AA.

XX AAV36744;

DT 27-SEP-1999 (first entry)

DE PGI splice variant allele protein sequence.

KW PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;  
 KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.

XX Homo sapiens.

XX WO9932644-A2.

XX 01-JUL-1999.

XX 22-DEC-1998; 98WO-IB02133.

XX 09-SEP-1998; 98US-0099658.

XX 22-DEC-1997; 97US-0996306.

XX (GEST ) GENSET.

XX

PI Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;

XX WPI: 1999-405178/34.

DR N-PSDB: AA200970.

XX

PT Use of a prostate cancer associated gene and biallelic markers

XX derived from it

XX

PS Claim 7; Page 250-251; 385pp; English.

XX

CC The invention relates to a mammalian PGI gene and protein, and a set of

CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are

CC used in a hybridisation assay, a sequencing assay, or in an

CC allele-specific amplification assay for determining the identity of a

CC nucleotide at a PGI-related biallelic marker. The methods can be used to

CC detect and to assess the risk of developing cancer or prostate cancer.

CC Early-stage diagnosis of prostate cancer relies on prostate specific

CC antigen (PSA) dosage. However, the effectiveness of this is limited due

CC to its inability to discriminate between malignant and non-malignant

CC affections of the organ. A need exists for both a reliable diagnostic

CC procedure which would enable early-stage diagnosis, and for preventative

CC and curative treatments of the disease. The PGI gene can be used for

CC detection of prostate cancer, and the risk of developing it in the

CC future, and can also be used to determine therapies for the disease.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

RESULT 12

AAV36740

ID AAV36740 standard; Protein; 228 AA.

XX AAV36740;

DT 27-SEP-1999 (first entry)

DE Truncated PGI protein sequence.

KW PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;  
 KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.

XX Homo sapiens.

XX WO9932644-A2.

XX 01-JUL-1999.

XX 22-DEC-1998; 98WO-IB02133.

XX 09-SEP-1998; 98US-0099658.

XX 22-DEC-1997; 97US-0996306.

XX

PA (GEST ) GENSET.  
XX  
PI Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;  
XX  
DR WPI: 1999-405178/34.  
XX N-PSDB: AA200926.  
PT Use of a prostate cancer associated gene and biallelic markers  
XX derived from it  
XX  
PS Claim 7; Page 208; 385pp; English.  
XX  
CC The invention relates to a mammalian PGI gene and protein, and a set of  
CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
CC used in a hybridisation assay, a sequencing assay, or in an  
CC allele-specific amplification assay for determining the identity of a  
CC nucleotide at a PGI-related biallelic marker. The methods can be used to  
CC detect and to assess the risk of developing cancer or prostate cancer.  
CC Early-stage diagnosis of prostate cancer relies on prostate specific  
CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
CC to its inability to discriminate between malignant and non-malignant  
CC affections of the organ. A need exists for both a reliable diagnostic  
CC procedure which would enable early-stage diagnosis, and for preventative  
CC and curative treatments of the disease. The PGI gene can be used for  
CC detection of prostate cancer, and the risk of developing it in the  
CC future, and can also be used to determine therapies for the disease.  
XX  
SQ Sequence 228 AA;

Query Match 64.6%; Score 228; DB 20; Length 228;  
Best Local Similarity 100.0%; Pred. No. 4.4e-224;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRVLLPSVVLGTAPTYVLANGVRLISAFLLPARYOALDDRLCYVQSMFLPEENTG 60  
Db 1 mrvllpsvvlgtaptyvlangvrlisafllparfyaldrrlcyvqsmvllfpeentg 60  
OY 61 VQILLYGDLPKNKENIITYLANHOSIVDVIADILAIROMALGHVRYVKEGLKMLPLTGC 120  
Db 61 vqilllygdlpknkenillylanhgstvdvdiadilaicrpalghvryvkeglkmlplygc 120  
OY 121 YFAOHGGIYVRSKAFNKEKRNKLOSYVDAGTPLYLVIPEEGTRYNPEQTKVLSASQAF 180  
Db 121 yfahggiiyvrsakfnkekrnklosyvda gtpmylvi fpeegtrynpeqtkvlisaqaf 180  
OY 181 AAOGLAVLKHVLPRIKATHVAFDCMKNYDAIYDVTVVVEGDDG 228  
Db 181 aagglavlkhlvprikathvafdcmknyda i ydv tvvvvegddg 228

RESULT 13  
AAV36745  
ID AAV36745 standard; Protein: 261 AA.  
XX  
AC AAV36745;  
XX  
DT 27-SEP-1999 (first entry)  
XX  
DE PGI splice variant allele protein sequence.

XX PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;  
KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.  
XX  
OS Homo sapiens.  
XX  
PN WO9932644-A2.  
XX  
PD 01-JUL-1999.  
XX  
PF 22-DEC-1998; 98WO-IB02133.  
XX  
PR 09-SEP-1998; 98US-0099658.

PR 22-DEC-1997; 97US-0996306.  
XX  
XX (GEST ) GENSET.  
XX  
PI Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;  
XX  
DR WPI: 1999-405178/34.  
XX N-PSDB: AA200971.  
PT Use of a prostate cancer associated gene and biallelic markers  
XX derived from it  
XX  
PS Claim 7; Page 251; 385pp; English.  
XX  
CC The invention relates to a mammalian PGI gene and protein, and a set of  
CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
CC used in a hybridisation assay, a sequencing assay, or in an  
CC allele-specific amplification assay for determining the identity of a  
CC nucleotide at a PGI-related biallelic marker. The methods can be used to  
CC detect and to assess the risk of developing cancer or prostate cancer.  
CC Early-stage diagnosis of prostate cancer relies on prostate specific  
CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
CC to its inability to discriminate between malignant and non-malignant  
CC affections of the organ. A need exists for both a reliable diagnostic  
CC procedure which would enable early-stage diagnosis, and for preventative  
CC and curative treatments of the disease. The PGI gene can be used for  
CC detection of prostate cancer, and the risk of developing it in the  
CC future, and can also be used to determine therapies for the disease.  
XX  
SQ Sequence 261 AA;

Query Match 56.4%; Score 199; DB 20; Length 261;  
Best Local Similarity 100.0%; Pred. No. 1.9e-194;  
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 155 MYLVIPEEGTRYNPEQTKVLSASQAFPAQRLAVLKHVLPRIKATHVAFDCMKNYDAI 214  
Db 63 mylvi fpeegtrynpeqtkvlisaqafaqr glavlkhvlprikathvafdcmknyda i 122  
OY 215 YDVTVVVEGKDDGGRRESPTMEFLCECPKIHIDRIDKVPPEOEHMRRLHERF 274  
Db 123 ydvtvvvegkddggrresptmeflckecpk ihidridkvppeoehmr r lherf 182  
OY 275 EIKDKMLIEFESDPERRKRPFGKSVNSKLSIKRTIPSMILSGLTAGMIMTDAGRKLX 334  
Db 183 eikdkmliefesdperrkrfp gksvnsksl sikt r tpsmilsgltagmimtdagr kly 242  
OY 335 VNTWITGTLGCLWVTIKA 353  
Db 243 vntwlygtllgclwvtika 261

RESULT 14  
AAV36753  
ID AAV36753 standard; Protein: 185 AA.  
XX  
AC AAV36753;  
XX  
DT 27-SEP-1999 (first entry)  
XX  
DE PGI splice variant allele protein sequence.

XX PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;  
KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.  
XX  
OS Homo sapiens.  
XX  
PN WO9932644-A2.  
XX  
PD 01-JUL-1999.  
XX  
PF 22-DEC-1998; 98WO-IB02133.



XX 09-SEP-1998; 98US-0099658.  
PR 22-DEC-1997; 97US-0996306.  
XX  
XX (GEST ) GENSET.  
PI Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;  
XX WPI; 1999-405178/34.  
DR N-PSDB; AA200979.  
XX  
XX Use of a prostate cancer associated gene and biallelic markers  
PT derived from it  
XX  
XX Claim 7; Page 255; 385pp; English.  
XX  
XX The invention relates to a mammalian PGI gene and protein, and a set of  
CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
CC used in a hybridisation assay, a sequencing assay, or in an  
CC allele-specific amplification assay for determining the identity of a  
CC nucleotide at a PGI-related biallelic marker. The methods can be used to  
CC detect and to assess the risk of developing cancer or prostate cancer.  
CC Early-stage diagnosis of prostate cancer relies on prostate specific  
CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
CC to its inability to discriminate between malignant and non-malignant  
CC affections of the organ. A need exists for both a reliable diagnostic  
CC procedure which would enable early-stage diagnosis, and for preventative  
CC and curative treatments of the disease. The PGI gene can be used for  
CC detection of prostate cancer, and the risk of developing it in the  
CC future, and can also be used to determine therapies for the disease.  
XX  
XX Sequence 185 AA;

Query Match 52.4%; Score 185; DB 20; Length 185;

Best Local Similarity 100.0%; Pred. No. 2.6e-180;

Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRLTSPSVLLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQSWLFEFFENYTG 60  
DB 1 MRLTSPSVLLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQSWLFEFFENYTG 60  
QY 61 VQILLTGDEPKKKNENTITLANHSTVDWIVADILAIQNALGHVRYLKGLKWLPLYGC 120  
DB 61 VQILLTGDEPKKKNENTITLANHSTVDWIVADILAIQNALGHVRYLKGLKWLPLYGC 120  
QY 121 YFAOHGSIYVKRSKAKFNEKEMRNKLSQSYVDAGTPMYLVIPPEGRVNPEDTKVLSASQAF 180  
DB 121 YFAOHGSIYVKRSKAKFNEKEMRNKLSQSYVDAGTPMYLVIPPEGRVNPEDTKVLSASQAF 180  
QY 181 AAORG 185  
DB 181 AAORG 185

RESULT 15

ID AA36751  
ID AA36751 standard; Protein; 315 AA.

AC AA36751;

DT 27-SEP-1999 (first entry)

DE PGI splice variant allele protein sequence.

XX PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;  
KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.  
XX

OS Homo sapiens.

XX WO932644-A2.

PD 01-JUL-1999.

XX 22-DEC-1998; 98WO-IB02133.  
XX

PR 09-SEP-1998; 98US-0099658.  
PR 22-DEC-1997; 97US-0996306.  
XX  
XX (GEST ) GENSET.

PI Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;

XX WPI; 1999-405178/34.

DR N-PSDB; AA200977.

XX Use of a prostate cancer associated gene and biallelic markers  
PT derived from it  
XX

XX Claim 7; Page 253-254; 385pp; English.

XX  
XX The invention relates to a mammalian PGI gene and protein, and a set of  
CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
CC used in a hybridisation assay, a sequencing assay, or in an  
CC allele-specific amplification assay for determining the identity of a  
CC nucleotide at a PGI-related biallelic marker. The methods can be used to  
CC detect and to assess the risk of developing cancer or prostate cancer.  
CC Early-stage diagnosis of prostate cancer relies on prostate specific  
CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
CC to its inability to discriminate between malignant and non-malignant  
CC affections of the organ. A need exists for both a reliable diagnostic  
CC procedure which would enable early-stage diagnosis, and for preventative  
CC and curative treatments of the disease. The PGI gene can be used for  
CC detection of prostate cancer, and the risk of developing it in the  
CC future, and can also be used to determine therapies for the disease.  
XX  
XX Sequence 315 AA;

Query Match 52.4%; Score 185; DB 20; Length 315;

Best Local Similarity 100.0%; Pred. No. 4.3e-180;

Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRLTSPSVLLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQSWLFEFFENYTG 60  
DB 1 MRLTSPSVLLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQSWLFEFFENYTG 60  
QY 61 VQILLTGDEPKKKNENTITLANHSTVDWIVADILAIQNALGHVRYLKGLKWLPLYGC 120  
DB 61 VQILLTGDEPKKKNENTITLANHSTVDWIVADILAIQNALGHVRYLKGLKWLPLYGC 120  
QY 121 YFAOHGSIYVKRSKAKFNEKEMRNKLSQSYVDAGTPMYLVIPPEGRVNPEDTKVLSASQAF 180  
DB 121 YFAOHGSIYVKRSKAKFNEKEMRNKLSQSYVDAGTPMYLVIPPEGRVNPEDTKVLSASQAF 180  
QY 181 AAORG 185  
DB 181 AAORG 185

Search completed: August 28, 2002, 11:18:44  
Job time: 582 sec

...

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:19:27 ; Search time 67.07 Seconds  
(without alignments)  
910.500 Million cell updates/sec

Title: US-09-853-526-4  
Perfect score: 353  
Sequence: 1 MRYLPSVLLGTAPTYLA.....YNTWVYGLGLMTWTFKA 353

Scoring table:  
OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 562222 seqs, 172994929 residues

Word size : 0

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

SPREMBL\_19:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	353	100.0	364	4 Q9B0G4	Q9B0G4 homo sapien
2	252	71.4	353	4 Q9NUQ2	Q9NUQ2 homo sapien
3	32	9.1	354	11 Q9D1E8	Q9D1E8 mus musculu
4	9	2.5	264	16 Q9POE7	Q9POE7 ureaplasma
5	9	2.5	386	5 Q9SRI2	Q9SRI2 caenorhabdi
6	8	2.3	115	16 Q9B8E7	Q9B8E7 rhizobium l
7	8	2.3	236	16 Q9A675	Q9A675 caulobacter
8	8	2.3	262	16 Q84743	Q84743 chlamydia t
9	8	2.3	282	16 Q83380	Q83380 treponema p
10	8	2.3	411	16 Q9A4V3	Q9A4V3 caulobacter
11	8	2.3	430	2 Q54517	Q54517 actinobact
12	8	2.3	454	3 Q9P5W2	Q9P5W2 neutrospora
13	8	2.3	458	16 Q9KFI0	Q9KFI0 bacillus ha
14	8	2.3	459	16 Q31503	Q31503 bacillus su
15	8	2.3	483	10 Q49927	Q49927 plasmu sativ
16	8	2.3	555	5 Q961R9	Q961R9 drosophila

## ALIGNMENTS

17	8	2.3	557	5 Q9VEX8	Q9VEX8 drosophila
18	8	2.3	1820	5 Q9VVG0	Q9VVG0 drosophila
19	7	2.0	54	6 Q28219	Q28219 cercopithec
20	7	2.0	78	16 Q98JCO	Q98JCO rhizobium l
21	7	2.0	126	11 Q62370	Q62370 mus musculu
22	7	2.0	132	11 Q9D581	Q9D581 mus musculu
23	7	2.0	137	12 Q55581	Q55581 leucania se
24	7	2.0	141	6 Q28218	Q28218 cercopithec
25	7	2.0	143	16 Q98RB9	Q98RB9 mycoplasma
26	7	2.0	148	2 Q9R25	Q9R25 streptomyce
27	7	2.0	148	16 Q9CH79	Q9CH79 lactococcus
28	7	2.0	150	2 Q9AGV0	Q9AGV0 bruceella ab
29	7	2.0	162	10 Q9LPP0	Q9LPP0 arabidopsis
30	7	2.0	163	2 Q9EX02	Q9EX02 streptomyce
31	7	2.0	175	2 Q9EX12	Q9EX12 klebsiella
32	7	2.0	184	11 Q9CWC3	Q9CWC3 mus musculu
33	7	2.0	195	16 Q9KCD7	Q9KCD7 bacillus ha
34	7	2.0	196	10 Q9LHL8	Q9LHL8 arabidopsis
35	7	2.0	209	16 Q910U7	Q910U7 pseudomonas
36	7	2.0	209	16 Q9CJKA	Q9CJKA pasteurella
37	7	2.0	211	16 Q67841	Q67841 aquifex aeo
38	7	2.0	220	2 Q9EW77	Q9EW77 streptomyce
39	7	2.0	225	16 Q97NE9	Q97NE9 streptococ
40	7	2.0	230	16 Q9PH25	Q9PH25 campylobact
41	7	2.0	234	2 Q32330	Q32330 clostridium
42	7	2.0	240	12 Q9WHT2	Q9WHT2 thailand to
43	7	2.0	240	16 P73054	P73054 synecocyst
44	7	2.0	241	16 Q97KFA	Q97KFA clostridium
45	7	2.0	247	16 Q9XDL6	Q9XDL6 neisseria m

## RESULT 1

ID	Q9B0G4	PRELIMINARY;	PRT;	364 AA.
AC	Q9B0G4:			
DT	01-JUN-2001 (TREMBLrel. 17, Created)			
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)			
DE	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	HYPOTHEICAL 42.1 KDA PROTEIN (LYSOPHOSPHATIDIC ACID			
DE	ACYLTRANSFERASE-EPSILON) (EC 2.3.1.51).			
GN	DKEZP761C22 OR LPAAT-E.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RC	SEQUENCE FROM N.A.			
RC	TISSUE=AMYGDAIA;			
RX	MDLINE=21154917; PubMed=11230166;			
RA	Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,			
RA	Ansoerge W., Boecher M., Bloeker H., Bauersachs S., Blum H.,			
RA	Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,			
RA	Mewes H.W., Ottenwelder B., Obermaier B., Tampe J., Heubner D.,			
RA	Wambutt R., Korn B., Klein M., Poustka A.,			
RT	Towards a Catalog of Human Genes and Proteins: Sequencing and			
RT	Analysis of 500 Novel Complete Protein Coding Human cDNAs."			
RL	Genome Res. 11:422-435(2001).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Leung D.W.;			
RT	"Cloning and expression of LPAAT-epsilon.";			
RT	Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AL136587; CAB66522.1; -			
DR	EMBL; AF375789; AAK54809.1; -			
DR	InterPro; IPR002123; Acyltransferase.			
DR	Pfam; PF01553; Acyltransferase; 1.			
KW	Hypothetical protein; transferase; Acyltransferase.			
SO	SEQUENCE 364 AA; 42072 MW; 90A0F87FC7C78081 CRC64;			

Query Match 100.0%; Score 353; DB 4; Length 364;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTAPTYVLAWGVRLLSAFLPARFYQALDRLCYVQSMVLEFFENTG 60  
 DB 12 MRYLLPSVLLGTAPTYVLAWGVRLLSAFLPARFYQALDRLCYVQSMVLEFFENTG 71  
 QY 61 VOILLYGDLPKKNENIYLANHGSTVDWIVADILAIRONALGHRYVYKEGLKMLPLXGC 120  
 DB 72 VOILLYGDLPKKNENIYLANHGSTVDWIVADILAIRONALGHRYVYKEGLKMLPLXGC 131  
 QY 121 YFAOHGGIYVRSKAFENEKERNKLSQSYVDAGTPMYLVIPEEGTRYNEDEQTKVLSASQAF 180  
 DB 132 YFAOHGGIYVRSKAFENEKERNKLSQSYVDAGTPMYLVIPEEGTRYNEDEQTKVLSASQAF 191  
 QY 181 AAOHGLAVLKHLVLPRIKATHVAFDCMKNYIDATYDVTVVVEGDDGQRESPTMTPEFL 240  
 DB 192 AAOHGLAVLKHLVLPRIKATHVAFDCMKNYIDATYDVTVVVEGDDGQRESPTMTPEFL 251  
 QY 241 CKECPKIHIDRIDKDVPEEOEHMRRLHEREFEIKDKMLIEFYESPDPERRRRFPCKS 300  
 DB 252 CKECPKIHIDRIDKDVPEEOEHMRRLHEREFEIKDKMLIEFYESPDPERRRRFPCKS 311  
 QY 301 VNSKLSIKTKLPSMLISGLTAGMLMDAGRKLYVNTWYIGTLLGCLMWTIKA 353  
 DB 312 VNSKLSIKTKLPSMLISGLTAGMLMDAGRKLYVNTWYIGTLLGCLMWTIKA 364

RESULT 2  
 Q9NUO2 PRELIMINARY; PRT; 353 AA.  
 ID Q9NUO2  
 AC Q9NUO2;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE CDNA FLJ11210 FTS, CLONE PLACE1007954.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=PLACENTA;  
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
 RA Magatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuno Y.,  
 RA Niimura Y., Iwayanagi T.;  
 RT "NEDO human cDNA sequencing project."  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AK002072; BAA92069.1; -;  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PFO1553; Acyltransferase; 1.  
 SQ SEQUENCE 353 AA; 40799 MW; 0ABAI4A2361F1B14 CRC64;

Query Match 71.4%; Score 252; DB 4; Length 353;  
 Best Local Similarity 99.7%; Pred. No. 3; Se-252;  
 Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTAPTYVLAWGVRLLSAFLPARFYQALDRLCYVQSMVLEFFENTG 60  
 DB 1 MRYLLPSVLLGTAPTYVLAWGVRLLSAFLPARFYQALDRLCYVQSMVLEFFENTG 60  
 QY 61 VOILLYGDLPKKNENIYLANHGSTVDWIVADILAIRONALGHRYVYKEGLKMLPLXGC 120  
 DB 61 VOILLYGDLPKKNENIYLANHGSTVDWIVADILAIRONALGHRYVYKEGLKMLPLXGC 120  
 QY 121 YFAOHGGIYVRSKAFENEKERNKLSQSYVDAGTPMYLVIPEEGTRYNEDEQTKVLSASQAF 180  
 DB 121 YFAOHGGIYVRSKAFENEKERNKLSQSYVDAGTPMYLVIPEEGTRYNEDEQTKVLSASQAF 180

DB 121 YFAOHGGIYVRSKAFENEKERNKLSQSYVDAGTPMYLVIPEEGTRYNEDEQTKVLSASQAF 180  
 QY 181 AAOHGLAVLKHLVLPRIKATHVAFDCMKNYIDATYDVTVVVEGDDGQRESPTMTPEFL 240  
 DB 181 AAOHGLAVLKHLVLPRIKATHVAFDCMKNYIDATYDVTVVVEGDDGQRESPTMTPEFL 240  
 QY 241 CKECPKIHIDRIDKDVPEEOEHMRRLHEREFEIKDKMLIEFYESPDPERRRRFPCKS 300  
 DB 241 CKECPKIHIDRIDKDVPEEOEHMRRLHEREFEIKDKMLIEFYESPDPERRRRFPCKS 300  
 QY 301 VNSKLSIKTKLPSMLISGLTAGMLMDAGRKLYVNTWYIGTLLGCLMWTIKA 353  
 DB 301 VNSKLSIKTKLPSMLISGLTAGMLMDAGRKLYVNTWYIGTLLGCLMWTIKA 353

RESULT 3  
 Q9DIE8 PRELIMINARY; PRT; 354 AA.  
 ID Q9DIE8  
 AC Q9DIE8;  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE 1110013A05RIK PROTEIN.  
 GN 1110013A05RIK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=EMBRYO;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arikawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
 RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Mashio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai J.H., Kohsaki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 DR EMBL: AK003649; BAB22915.1; -;  
 DR MGD: MGI:1915880; 1110013A05RIK.  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PFO1553; Acyltransferase; 1.  
 SQ SEQUENCE 354 AA; 40943 MW; 447EEF924B91E800 CRC64;

Query Match 9.1%; Score 32; DB 11; Length 354;  
 Best Local Similarity 100.0%; Pred. No. 1; 9e-24;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 174 LSASQAFPAOGLAVLKHLVLPRIKATHVARD 205  
 DB 174 LSASQAFPAOGLAVLKHLVLPRIKATHVARD 205  
 RESULT 4  
 Q9POE7 PRELIMINARY; PRT; 264 AA.  
 ID Q9POE7;  
 AC Q9POE7;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)

DT 01-OCT-2000 (TREMblrel. 15, last sequence update)  
 DT 01-JUN-2001 (TREMblrel. 17, last annotation update)  
 DE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE.  
 GN PLSC OR UU344.  
 OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
 OC Mycoplasmataceae; Ureaplasma.  
 OX NCBI\_TaxID=134821;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SEROVAR 3;  
 RX MEDLINE=20500219; PubMed=11048724;  
 RA Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,  
 RA Cassell G.H.;  
 RT "The complete sequence of the mucosal pathogen Ureaplasma  
 RT urealyticum";  
 RT Nature 407:757-762(2000).  
 DR EMBL; AF002131; AAF30753.1; -;  
 DR InterPro; IPR002123; Acyltransferase.  
 DR Pfam; PF01553; Acyltransferase; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 264 AA; 30687 MW; DBC5727A07C7AF7 CRC64;

Query Match 2.5%; Score 9; DB 16; Length 264;  
 Best Local Similarity 100.0%; Pred. No. 0.97;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 157 LVIFPECTR 165  
 DB 156 LVIFPECTR 164

RESULT 5  
 Q95R12 PRELIMINARY; PRT; 386 AA.  
 AC Q95R12;  
 DT 01-DEC-2001 (TREMblrel. 19, Created)  
 DT 01-DEC-2001 (TREMblrel. 19, last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)  
 DE HYPOTHETICAL 44.0 KDA PROTEIN.  
 GN F28B3.9.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA None;  
 RT "Genome sequence of the nematode C. elegans: a platform for  
 RT investigating biology. The C. elegans Sequencing Consortium.";  
 RT Science 282:2012-2018(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Geisel C., Kramer J., Smith A.;  
 RT "The sequence of C. elegans cosmid F28B3.";  
 RT submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Waterston R.;  
 RT "Direct Submission";  
 RT Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF003136; AAK93853.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 386 AA; 43978 MW; A1E3020C201452A6 CRC64;

Query Match 2.5%; Score 9; DB 5; Length 386;  
 Best Local Similarity 100.0%; Pred. No. 1.3;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 211 LDAIYDVT 219  
 DB 230 LDAIYDVT 238

RESULT 6  
 Q98E87 PRELIMINARY; PRT; 115 AA.  
 ID Q98E87;  
 AC Q98E87;  
 DT 01-OCT-2001 (TREMblrel. 18, Created)  
 DT 01-OCT-2001 (TREMblrel. 18, last sequence update)  
 DE MLL4358 PROTEIN.  
 GN MLL4358.  
 OS Rhizobium loti (Mesorhizobium loti).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Phyllobacteriaceae; Mesorhizobium.  
 OX NCBI\_TaxID=381;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-MAFE30309;  
 RX MEDLINE=21082930; PubMed=11214968;  
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
 RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,  
 RA Kouchizuki Y., Nakayama S., Nakazaki N., Shimoto S., Sugimoto M.,  
 RA Takeuchi C., Yamada M., Tabata S.;  
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
 RT Mesorhizobium loti.";  
 RT DNA Res. 7:331-338(2000).  
 DR EMBL; AP003004; BAB51033.1; -;  
 KW Complete proteome.  
 SQ SEQUENCE 115 AA; 12335 MW; 6AF43D6451F42BE8 CRC64;

Query Match 2.3%; Score 8; DB 16; Length 115;  
 Best Local Similarity 100.0%; Pred. No. 5.1;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 174 LSASQAF 181  
 DB 31 LSASQAF 38

RESULT 7  
 Q9A675 PRELIMINARY; PRT; 236 AA.  
 ID Q9A675;  
 AC Q9A675;  
 DT 01-JUN-2001 (TREMblrel. 17, Created)  
 DT 01-JUN-2001 (TREMblrel. 17, last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)  
 DE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE.  
 GN CC2219.  
 OS Caulobacter crescentus.  
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
 OC Caulobacter.  
 OX NCBI\_TaxID=69394;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 19089 / CB15;  
 RX MEDLINE=21173698; PubMed=11259647;  
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
 RA Pollock A., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
 RA DeBoy R.T., Dosson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,  
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,  
 RA Uterback T., Tran K., Wolf A., Yamathavan J., Ernolova M., White O.,  
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
 RT "Complete genome sequence of Caulobacter crescentus.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 DR EMBL; AE005893; AAK24190.1; -;

DR TIGR; CC2219; -  
 DR InterPro; IPR002123; Acyltransferase.  
 DR Pfam; PF01553; Acyltransferase; 1.  
 KW Transferrase; Acyltransferase; Complete proteome.  
 SO SEQUENCE 236 AA; 26382 MW; 85694D3966BCA1A0 CRC64;

Query Match 2.3%; Score 8; DB 16; Length 236;  
 Best Local Similarity 100.0%; Pred. No. 9.5;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 158 VIFPEGTR 165  
 Db 146 VIFPEGTR 153

RESULT 8  
 084743 PRELIMINARY; PRT; 262 AA.  
 AC 084743;  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE METAL DEPENDENT HYDROLASE.  
 GN YXCU OR CT738.  
 OS Chlamydia trachomatis.  
 OC Bacteria; Chlamydiales; Chlamydiales; Chlamydia.  
 OX NCBI\_Taxid=813;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=D/UW-3/CX;  
 RX MEDLINE=99000809; PubMed=9784136;  
 RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,  
 RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,  
 RA Davis R.W.;  
 RT "Genome sequence of an obligate intracellular pathogen of humans:  
 RT Chlamydia trachomatis.";  
 RL Science 282:754-759(1998).  
 DR EMBL; AE001344; AAC68333.1; -.  
 KW Hydrolase; Complete proteome.  
 SO SEQUENCE 262 AA; 29082 MW; E36A6298A23CEFBF CRC64;

Query Match 2.3%; Score 8; DB 16; Length 262;  
 Best Local Similarity 100.0%; Pred. No. 10;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 191 HVLPRPRK 198  
 Db 202 HVLPRPRK 209

RESULT 9  
 083380 PRELIMINARY; PRT; 282 AA.  
 AC 083380;  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE LYSOPHOSPHATIDIC ACID ACYLTRANSFERASE, PUTATIVE.  
 GN TP0361.  
 OS Treponema pallidum.  
 OC Bacteria; Spirochaetales; Spirochaetales; Treponema.  
 OX NCBI\_Taxid=160;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NICHOLS;  
 RX MEDLINE=98332770; PubMed=9665876;  
 RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,  
 RA Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,  
 RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,  
 RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,  
 RA McDonald L., Artlich F., Bowman C., Colton M.D., Fujii C., Garland S.,

RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,  
 RA Venter J.C.;  
 RT "Complete genome sequence of Treponema pallidum, the syphilis  
 RT spirochete.";  
 RL Science 281:375-388(1998).  
 DR EMBL; AE001215; AAC65346.1; -.  
 DR TIGR; TP0361; -.  
 DR InterPro; IPR002123; Acyltransferase.  
 DR Pfam; PF01553; Acyltransferase; 1.  
 KW Transferrase; Acyltransferase; Complete proteome.  
 SO SEQUENCE 282 AA; 31700 MW; DEF6903CD6C61AD CRC64;

Query Match 2.3%; Score 8; DB 16; Length 282;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 158 VIFPEGTR 165  
 Db 170 VIFPEGTR 177

RESULT 10  
 094AV3 PRELIMINARY; PRT; 411 AA.  
 AC 094AV3;  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE METAL ION EFFLUX MEMBRANE FUSION PROTEIN FAMILY.  
 GN CC2722.  
 OS Caulobacter crescentus.  
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
 OC Caulobacter.  
 OX NCBI\_Taxid=69394;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 19089 / CB15;  
 RX MEDLINE=21173698; PubMed=11259647;  
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadtke N.D., Ely B.,  
 RA Deboy R.T., Dodson R.D., Durkin A.S., Gwin M.L., Haft D.H.,  
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,  
 RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,  
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
 RT "Complete genome sequence of Caulobacter crescentus.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 DR EMBL; AE005938; AAK24687.1; -.  
 DR TIGR; CC2722; -.  
 DR InterPro; IPR002215; H1YD.  
 DR Pfam; PF00529; H1YD; 1.  
 KW Complete proteome.  
 SO SEQUENCE 411 AA; 41818 MW; 1BA7849B9BCF22E CRC64;

Query Match 2.3%; Score 8; DB 16; Length 411;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 316 ILSGLTAG 323  
 Db 380 ILSGLTAG 387

RESULT 11  
 054517 PRELIMINARY; PRT; 430 AA.  
 AC 054517;  
 DT 01-JUN-1998 (TREMBlrel. 06, Created)  
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)  
 DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)  
 DE 3-DEOXY-D-MANNNO-2-OCTULOSONATE TRANSFERASE.

GN WAAA.  
OS Acinetobacter baumannii.  
OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;  
OC Acinetobacter.  
OX NCBL\_TaxID=470;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 19606, AND ATCC 15308;  
RA Bode C.E., Brabetz W., Brade H.;  
RL Submitted (JAN1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: Z56926; CAB09652.1; -  
DR EMBL: Z56925; CAB09651.1; -  
KW Transferase.  
SQ SEQUENCE 430 AA; 49070 MW; 695293748185D272 CRC64;

Query Match 2.3%; Score 8; DB 2; Length 430;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 318 SGLTAGML 325  
Db 171 SGLTAGML 178

RESULT 12  
O9P5W2 PRELIMINARY; PRT; 454 AA.  
AC O9P5W2;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)  
DE 01-DEC-2001 (TREMBLrel. 19, last annotation update)  
DE RELATED TO GUANINE DEAMINASE.  
GN B3E4.200.  
OS Neurospora crassa.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Sordariales; Sordariaceae; Neurospora.  
OX NCBL\_TaxID=5141;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Schulte U., Align V., Hoheisel J., Brandt P., Fartmann B., Holland R.,  
RA Nyakatura G., Mewes H.W., Mannhaupt G.;  
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA German Neurospora genome project;  
RA Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
RL EMBL: AL355931; CAB91408.2; -  
SQ SEQUENCE 454 AA; 49778 MW; 8C1EC14BAECA421D CRC64;

Query Match 2.3%; Score 8; DB 3; Length 454;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 189 LKHVTPR 196  
Db 208 LKHVTPR 215

RESULT 13  
O9KF10 PRELIMINARY; PRT; 458 AA.  
AC O9KF10;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)  
DE 01-DEC-2001 (TREMBLrel. 19, last annotation update)  
DE RNA METHYLTRANSFERASE.  
GN BH0687.  
OS Bacillus halodurans.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Bacillus.  
OX NCBL\_TaxID=86665;

RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-C-125 / JCM 9153;  
RX MEDLINE=20512562; PubMed=11058132;  
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,  
RA Horikoshi K.;  
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus  
RT halodurans and genomic sequence comparison with Bacillus subtilis.";  
RL Nucleic Acids Res. 28:4317-4331(2000).  
DR EMBL: AP001509; BAB04406.1; -  
DR InterPro: IPR000051; SAM\_bind.  
DR InterPro: IPR001566; TRMA\_1.  
DR PROSITE: PS01230; TRMA\_1; UNKNOWN\_1.  
KW Transferase; Methyltransferase; Complete proteome.  
SQ SEQUENCE 458 AA; 51555 MW; 3D33A651E8B91240 CRC64;

Query Match 2.3%; Score 8; DB 16; Length 458;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 167 NPEQTKVL 174  
Db 291 NPEQTKVL 298

RESULT 14  
O31503 PRELIMINARY; PRT; 459 AA.  
AC O31503;  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, last sequence update)  
DT 01-OCT-2001 (TREMBLrel. 18, last annotation update)  
DE YEFA PROTEIN.  
GN YEFA.  
OS Bacillus subtilis.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Bacillus.  
OX NCBL\_TaxID=1423;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-168;  
RX MEDLINE=98044033; PubMed=9384377;  
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
RA Azevedo V., Bertero M.G., Bessières P., Bolotin A., Borchert S.,  
RA Borriis R., Boursier L., Brans A., Brun M., Brignell S.C., Bron S.,  
RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,  
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,  
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
RA Fritz C., Fujita M., Fujita Y., Funa S., Gallizzi A., Galleron N.,  
RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,  
RA Giuseppe G., Guy B.J., Haga K., Haeche J., Harwood C.R., Henaut A.,  
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
RA Joris B., Karamata D., Kasahara Y., Klier-Bianchard M., Klein C.,  
RA Kobayashi Y., Koelter P., Koningsstein G., Krogh S., Kumano M.,  
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,  
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,  
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,  
RA Sorokin A., Taccioni E., Takagi T., Takahashi H., Takemaru K.,  
RA Takeuchi M., Tamakoshi A., Tanaka T., Tempstra P., Tognoni A.,  
RA Tosato V., Uchiyama S., Vandenbol M., Vamier F., Vassarotti A.,  
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,  
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,  
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,  
RT "The complete genome sequence of the gram-positive bacterium Bacillus

RT sub111s.":  
 RL Mature 390:249-256(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-168:  
 RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.:  
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z99107; CAB12493.1; -.  
 DR InterPro: IPR002792; DUF90.  
 DR InterPro: IPR000051; SAM\_bind.  
 DR InterPro: IPR001565; TRMA\_1.  
 DR Pfam: PF01938; TRAM; 1.  
 DR PROSITE; PS01230; TRMA\_1; UNKNOWN\_1.  
 DR PROSITE; PS01231; TRMA\_2; 1.  
 KM Complete proteome.  
 SQ SEQUENCE 459 AA; 51819 MW; 35E99547779421D5 CRC64;

Query Match 2.38; Score 8; DB 16; Length 459;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Caps 0;

OY 167 NPEQTKVL 174  
 |||||  
 DB 292 NPEQTKVL 299

RESULT 15  
 049927 PRELIMINARY; PRT; 483 AA.  
 AC 049927;  
 DT 01-JUN-1998 (TREMBLrel. 06, Created)  
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE P54. PROTEIN.  
 GN P54.  
 OS Pisum sativum (Garden pea).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.  
 OC NCBI\_TaxID=3888;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Castillo J., Marquez J.A., Franco L., Ballestar E., Rodrigo M.I.;  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Y11207; CAA72090.1; -.  
 DR HSSP: P50477; ICAU.  
 DR InterPro: IPR001113; Seedstore\_7s.  
 DR Pfam: PF00546; Seedstore\_7s; 1.  
 DR Pfam: PF02808; Seedstore\_7s\_C; 1.  
 SQ SEQUENCE 483 AA; 54662 MW; 8127BDAAA0178F3D CRC64;

Query Match 2.38; Score 8; DB 10; Length 483;  
 Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 151 AGTPMYLV 158  
 |||||  
 DB 184 AGTPMYLV 191

Search completed: August 28, 2002, 11:28:33  
 Job time: 546 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 10:59:36 ; Search time 75.55 Seconds  
(without alignments)  
518.981 Million cell updates/sec

Title: US-09-853-526-4

Perfect score: 1871

Sequence: 1 MRYLLPSVILGTAFTYVLA.....YVNTWIVGTLGLMTWTKA 353

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing:

Maximum Match 0%

Listing first 45 summaries

Database : A.Geneseq\_032802.\*

- 1: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT.\*
- 2: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT.\*
- 3: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT.\*
- 4: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT.\*
- 5: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT.\*
- 6: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT.\*
- 7: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT.\*
- 8: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT.\*
- 9: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT.\*
- 10: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT.\*
- 11: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT.\*
- 12: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT.\*
- 13: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT.\*
- 14: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT.\*
- 15: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT.\*
- 16: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT.\*
- 17: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT.\*
- 18: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT.\*
- 19: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT.\*
- 20: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT.\*
- 21: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT.\*
- 22: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1871	100.0	353	22	AAB85532	Human secreted pro
2	1871	100.0	353	22	AAB85552	Human secreted pro
3	1871	100.0	364	20	AA36729	Human PGI protein
4	1871	100.0	364	22	AA50128	Human acyltransfer
5	1871	100.0	364	22	AA000655	Human lysophosphat
6	1868	99.8	353	22	AA035955	Human protein sequ
7	1867	99.8	352	20	AA36728	Human PGI protein
8	1861	99.5	353	22	AA39591	Human polypeptide
9	1860	99.4	372	22	AA41377	Human polypeptide
10	1644	87.9	315	20	AA36751	PGI splice variant
11	1568.5	83.8	354	20	AA36741	Mouse PGI protein

12	1555.5	83.1	300	20	AA36752	PGI splice variant
13	1498	80.1	291	20	AA36744	PGI splice variant
14	1427	76.3	269	22	AAE09653	Human gene 15 enco
15	1325	70.8	261	20	AA36745	PGI splice variant
16	1203	64.3	228	20	AA36740	Truncated PGI prot
17	1182.5	63.2	238	20	AA36743	PGI splice variant
18	976	52.2	185	20	AA36753	PGI splice variant
19	820	43.8	182	20	AA36750	PGI splice variant
20	449	24.0	96	20	AA36749	PGI splice variant
21	401.5	21.5	378	21	AA36830	Arabidopsis thalia
22	396.5	21.2	352	21	AA36831	Arabidopsis thalia
23	393.5	21.0	341	21	AA36832	Arabidopsis thalia
24	390.5	20.9	375	21	AA36734	Arabidopsis thalia
25	390.5	20.9	375	21	AA36735	Arabidopsis thalia
26	387.5	20.7	351	21	AA36735	Arabidopsis thalia
27	387.5	20.7	375	21	AA36804	Arabidopsis thalia
28	387	20.7	257	22	ABG01360	Novel human diagno
29	387	20.7	980	22	ABG11977	Novel human diagno
30	384.5	20.6	351	21	AA36746	Arabidopsis thalia
31	374.5	20.0	343	21	AA36736	Arabidopsis thalia
32	374.5	20.0	343	21	AA36742	Arabidopsis thalia
33	332	17.7	77	20	AA36742	PGI splice variant
34	330	17.6	66	20	AA36748	PGI splice variant
35	329	17.6	68	20	AA36747	PGI splice variant
36	329	17.6	90	20	AA36746	PGI splice variant
37	328	17.5	374	22	AA000667	Maize lysophosphat
38	326.5	17.5	374	21	AA008479	Maize acid sequenc
39	326	17.4	374	15	AA36712	Maize 2-acylttransf
40	315.5	16.9	374	21	AA008477	Amino acid sequenc
41	311.5	16.6	387	21	AA36748	Jojoba acyltransfe
42	310	16.6	377	17	AA36749	Limonathes CDNA cl
43	309.5	16.5	294	21	AA36741	zee mays protein I
44	308.5	16.5	389	21	AA36741	Arabidopsis thalia
45	307.5	16.4	389	21	AA36741	Arabidopsis thalia

#### ALIGNMENTS

RESULT 1	
AAB85532	standard; protein; 353 AA.
ID	
AC	AAB85532;
XX	
DT	25-SEP-2001 (first entry)
XX	
DE	Human secreted protein (clone Id HUTH084).
XX	
KW	Secreted protein; immunosuppressive; antiarthritic; antirheumatic;
KW	antiproliferative; cyostatic; cardiatic; vasotropic; cerebroprotective;
KW	nootropic; neuroprotective; antibacterial; virucide; fungicide; human;
KW	ophthalmological; gene therapy.
OS	
XX	Homo sapiens.
XX	
PN	MO200155430-A1.
XX	
PD	02-AUG-2001.
XX	
PF	17-JAN-2001; 2001MO-US01431.
XX	
PR	31-JAN-2000; 2000US-0179065.
PR	04-FEB-2000; 2000US-0180628.
PR	12-SEP-2000; 2000US-0231968.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Rosen CA, Komatsoulis GA, Baker KP, Blise CE, Soppet DR, Olsen HS;
PI	Moore PA, Wei P, Ehner R, Duan DR, Shi Y, Choi GH, Fiselilla M;
PI	Ni J, Ruben SM, Barash SC;
XX	
DR	WPI; 2001-476220/51.

DR N-PSDB; AAH46942.  
XX 17 Isolated nucleic acid molecules encoding human secreted proteins,  
PT used to preventing, treating or ameliorating a medical condition  
XX  
PS Claim 11; Page 450-451; 482pp; English.  
XX  
XX The invention provides novel human secreted proteins and polynucleotides  
CC encoding them. The secreted proteins can be expressed by standard  
CC recombinant methodology. The secreted proteins and polynucleotides are  
CC used to prevent, treat or ameliorate a medical condition in e.g. humans,  
CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They can  
CC also be used in diagnosing a pathological condition. The antibodies to  
CC the proteins can also be used in alleviating symptoms associated with the  
CC disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme  
CC linked immunosorbent assays (ELISA). Disorders which are diagnosed or  
CC treated include autoimmune diseases e.g. rheumatoid arthritis,  
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,  
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
CC e.g. cerebral ischemia, angiodenesis, nervous system disorders e.g.  
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and  
CC ocular disorders e.g. corneal infection. The polypeptides can also be  
CC used to aid wound healing and epithelial cell proliferation, to prevent  
CC skin aging due to sunburn, to maintain organs before transplantation, for  
CC supporting cell culture of primary tissues, to regenerate tissues and in  
CC chemotaxis. The polypeptides can also be used as a food additive or  
CC preservative to increase or decrease storage capabilities. The present  
CC sequence represents a human secreted protein.  
XX  
SQ Sequence 353 AA:  
  
Query Match 100.0%; Score 1871; DB 22; Length 353;  
Best Local Similarity 100.0%; Pred. No. 2e-189;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MRYLLPSVVLGTAFTYVLAAGVWRLLSAFLPAPFYQALDDRLCYQSWLFFFEFNTG 60  
DB 1 mryllpsvvlgtapcyvlawgvrllsalfparfyqalddrlcyvqswmlfffeentyg 60  
  
QY 61 VOILLYGDLPRKNENITYLANHSTVDMIVADILAIROMALGHVRVYLKEGKLPLGYC 120  
DB 61 vqilllygdiprknkenillylanhgstvdvialaiqnalghvrlylkegkwpilygc 120  
  
QY 121 YFAOHGCIYVKSAAKNEKEMRNKLSGYVDAGTPMVLVIFPEGSTRYNPEQTVLSASQAF 180  
DB 121 yfaohgciyvksaaknekenrnklsqyvda gtpmvlvifpegst rypnpeqtkvlsasqaf 180  
  
QY 181 AAOGLAVLKHVLTPEIKTHAFDCMKNYLDAIVDVTYVYEGKDDGOGRESTPTMEFL 240  
DB 181 aagrglavlkhvltprtkthafdcmknyldaiydvtyvvegkddgqgrresptmlefl 240  
  
QY 241 CKECPKTHIHDIRDKRDVEDEHMRRLHERFEIKDKMLIEFYESPDERRRKRPFGKS 300  
DB 241 ckecpkthihdirdkrdvdeehmrmlherfeikdkmliefyespderrrkrrpfpgks 300  
  
QY 301 VNSKLSIKTKLPSMLISGLTAGMLMTDAGRKLVTNTWIVGTLLGCLMWTIKA 353  
DB 301 vnsklslktklpsmlisgltagmlmtdagrklvntwivgtllgclmwitika 353  
  
RESULT 2  
AAB85552  
ID AAB85552 standard; protein; 353 AA.  
AC AAB85552;  
XX  
XX  
DT 25-SEP-2001 (first entry)  
XX  
XX Human secreted protein (clone id HSLA181).  
XX  
XX Secreted protein; immunosuppressive; antitubercic; antirheumatic;  
KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;

KW neurotropic; neuroprotective; antibacterial; virucide; fungicide; human;  
KW ophthalmological; gene therapy.  
XX  
XX Homo sapiens.  
XX  
XX WO20015430-A1.  
XX  
XX 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US01431.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
XX 04-FEB-2000; 2000US-0180628.  
XX 12-SEP-2000; 2000US-0231968.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR, Olsen HS;  
PI Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH, Fiscella M;  
PI Ni J, Ruben SM, Barash SC;  
XX  
XX WPI: 2001-476220/51.  
DR N-PSDB; AAH46942.  
DR  
XX  
XX 17 Isolated nucleic acid molecules encoding human secreted proteins,  
PT used to preventing, treating or ameliorating a medical condition  
XX  
PS Claim 11; Page 465-466; 482pp; English.  
XX  
XX The invention provides novel human secreted proteins and polynucleotides  
CC encoding them. The secreted proteins can be expressed by standard  
CC recombinant methodology. The secreted proteins and polynucleotides are  
CC used to prevent, treat or ameliorate a medical condition in e.g. humans,  
CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They can  
CC also be used in diagnosing a pathological condition. The antibodies to  
CC the proteins can also be used in alleviating symptoms associated with the  
CC disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme  
CC linked immunosorbent assays (ELISA). Disorders which are diagnosed or  
CC treated include autoimmune diseases e.g. rheumatoid arthritis,  
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,  
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
CC e.g. cerebral ischemia, angiodenesis, nervous system disorders e.g.  
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and  
CC ocular disorders e.g. corneal infection. The polypeptides can also be  
CC used to aid wound healing and epithelial cell proliferation, to prevent  
CC skin aging due to sunburn, to maintain organs before transplantation, for  
CC supporting cell culture of primary tissues, to regenerate tissues and in  
CC chemotaxis. The polypeptides can also be used as a food additive or  
CC preservative to increase or decrease storage capabilities. The present  
CC sequence represents a human secreted protein.  
XX  
SQ Sequence 353 AA:  
  
Query Match 100.0%; Score 1871; DB 22; Length 353;  
Best Local Similarity 100.0%; Pred. No. 2e-189;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MRYLLPSVVLGTAFTYVLAAGVWRLLSAFLPAPFYQALDDRLCYQSWLFFFEFNTG 60  
DB 1 mryllpsvvlgtapcyvlawgvrllsalfparfyqalddrlcyvqswmlfffeentyg 60  
  
QY 61 VOILLYGDLPRKNENITYLANHSTVDMIVADILAIROMALGHVRVYLKEGKLPLGYC 120  
DB 61 vqilllygdiprknkenillylanhgstvdvialaiqnalghvrlylkegkwpilygc 120  
  
QY 121 YFAOHGCIYVKSAAKNEKEMRNKLSGYVDAGTPMVLVIFPEGSTRYNPEQTVLSASQAF 180  
DB 121 yfaohgciyvksaaknekenrnklsqyvda gtpmvlvifpegst rypnpeqtkvlsasqaf 180  
  
QY 181 AAOGLAVLKHVLTPEIKTHAFDCMKNYLDAIVDVTYVYEGKDDGOGRESTPTMEFL 240  
DB 181 aagrglavlkhvltprtkthafdcmknyldaiydvtyvvegkddgqgrresptmlefl 240

OY	241	CKECPKHIIHIDRIDKDVPEQEHMRWRLHEREIREIKDKMLIEFYESPDEBRKRPFGRK	300
Db	241	CECPKHIIHIDRIDKDVPEQEHMRWRLHEREIREIKDKMLIEFYESPDEBRKRPFGRK	300
OY	301	VNSKLSIKTKLPMSMLISGLTAGMLMTDAGRKILYNNVIYGTLLGCLMTVITKA	353
Db	301	VNSKLSIKTKLPMSMLISGLTAGMLMTDAGRKILYNNVIYGTLLGCLMTVITKA	353
RESULT	3		
ID	AA36729	standard; Protein; 364 AA.	
AC	AA36729;		
XX			
XX	27-SEP-1999	(first entry)	
XX			
DE	Human PGI protein sequence.		
XX			
KW	PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;		
XX	cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.		
XX			
OS	Homo sapiens.		
PN	MO9932644-A2.		
XX			
PD	01-JUL-1999.		
XX			
PF	22-DEC-1998; 98WO-1B02133.		
XX			
PR	09-SEP-1998; 98US-0099658.		
XX			
PR	22-DEC-1997; 97US-0996306.		
XX			
PA	(GEST ) GENSET.		
XX			
PI	Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;		
XX			
XX	WPI: 1999-405178/34.		
DR	N-PSDB; AA200870.		
XX			
PT	Use of a prostate cancer associated gene and biallelic markers		
XX	derived from it		
XX			
PS	Claim 7; Page 190-191; 385pp; English.		
XX			
CC	The invention relates to a mammalian PGI gene and protein, and a set of		
XX	PGI biallelic markers. The PGI polynucleotide and biallelic markers are		
XX	used in a hybridisation assay, a sequencing assay, or in an		
CC	allele-specific amplification assay for determining the identity of a		
XX	nucleotide at a PGI-related biallelic marker. The methods can be used to		
CC	detect and to assess the risk of developing cancer or prostate cancer.		
XX	Early-stage diagnosis of prostate cancer relies on prostate specific		
CC	antigen (PSA) dosage. However, the effectiveness of this is limited due		
XX	to its inability to discriminate between malignant and non-malignant		
CC	affections of the organ. A need exists for both a reliable diagnostic		
XX	procedure which would enable early-stage diagnosis, and for preventative		
CC	and curative treatments of the disease. The PGI gene can be used for		
XX	detection of prostate cancer, and the risk of developing it in the		
CC	future, and can also be used to determine therapies for the disease.		
XX			
SO	Sequence 364 AA;		
	Query Match	100.0%; Score 1871; DB 20; Length 364;	
	Best Local Similarity	100.0%; Pred. No. 2,1e-189;	
	Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps		
OY	1	MRYLPSVVLGTAPTVLAWGVWRLISATLPARFYALDDRLRCVQSVLFEFFENYTG	60
Db	12	mryllpsvvlgtaptvylawgvwrlisatlparfyalddrlrcvqsvlffennycg	71
OY	61	VQILLGDLPEKKNENIYYLANHOSVTWIVADILAIROMALGHVRYVLEKGLMLPYGC	120

Db	72	vglllygdipknkeniilylanhgstvdwivadlialrqnalghvrryilkeglkwlpygc	131
Qy	121	YFAQHGCIYVRSAKFNENKERNKLGQSYVDAGTPMYLVIEPEGTRVNEQOTKYLASQAF	180
Db	132	yfaqhgciyvrtsakfnekemnrnlqsyvdagtpmylvilfpegtrvneqtkylasaqaf	191
Qy	181	AAQRLAVLKVHLVPRIKATHVAFDCMKNTYDITVTVYEGKDDGGQRRSPRTMTEFL	240
Db	192	aagrglavlkvhlvprlkatvafdcmknyldaiydvltvyegddgqrrspmttefl	251
Qy	241	CKECKTHIHIDRIDKKVPEQDGHMRWMLHERREIKDKMLIEYESPDPERRRKRPFGKS	300
Db	252	ckeckthihidridkkvpeeqdghmrwmlherreikdkmlieyespdperrrkrfpgks	311
Qy	301	VNSKLSIKTKLPMSMLISGLTAGMLMTDAGCKRLVTNTWYITGLGLCLMTYIKA	353
Db	312	vnsklsiktklpmsmlisgltagmlmtdagckrlvntwiygtllgclwtyika	364
RESULT 4			
AA	50128		
ID	AA050128	standard; Protein; 364 AA.	
XX	AA050128;		
AC	AA050128;		
XX	21-DEC-2001 (first entry)		
DT	Human acyltransferase 27417.		
XX			
XX			
KW	Acyltransferase 27417; human; cell proliferation; cell migration;		
KW	cell differentiation; ovary cancer; brain cancer; colon cancer;		
KW	lung cancer; tumour; metastasis; sarcoma; carcinoma;		
KW	adenocarcinoma; antitumour; lipid metabolism; diagnosis; therapy.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FH	Peptide	1..49	
FT	Protein	/label= Signal_peptide	
FT	Domain	50..364	
FT	Domain	/label= Mature_protein	
FT	Domain	50..320	
FT	Domain	/label= "O-terminal_domain	
FT	Domain	321..337	
FT	Domain	/label= "Transmembrane_domain	
FT	Domain	338..364	
FT	Domain	/label= C-terminal_domain	
FT	Domain	71..363	
FT	Domain	/label= Acyltransferase_domain	
FT	Modified-site	68..71	
FT	Modified-site	/note= "Asn is N-glycosylated"	
FT	Modified-site	241..244	
FT	Modified-site	/note= "O-phosphorylated by cAMP- and cGMP-dependent protein kinase"	
FT	Modified-site	11..13	
FT	Modified-site	/note= "O-phosphorylated by protein kinase C"	
FT	Modified-site	144..146	
FT	Modified-site	/note= "O-phosphorylated by protein kinase C"	
FT	Modified-site	205..207	
FT	Modified-site	/note= "O-phosphorylated by protein kinase C"	
FT	Modified-site	317..319	
FT	Modified-site	/note= "O-phosphorylated by protein kinase C"	
FT	Modified-site	361..363	
FT	Modified-site	/note= "O-phosphorylated by protein kinase C"	
FT	Modified-site	95..98	
FT	Modified-site	/note= "O-phosphorylated by casein kinase II"	
FT	Modified-site	158..161	
FT	Modified-site	/note= "O-phosphorylated by casein kinase II"	
FT	Modified-site	246..249	
FT	Modified-site	/note= "O-phosphorylated by casein kinase II"	
FT	Modified-site	82..89	
FT	Modified-site	/note= "O-phosphorylated by tyrosine kinase"	

FT	Modified-site	219..226	/note= "O-phosphorylated by tyrosine kinase"
FT	Modified-site	23..28	/note= "N-myristoylated"
FT	Modified-site	130..135	/note= "N-myristoylated"
FT	Modified-site	330..335	/note= "N-myristoylated"
FT	Modified-site	352..357	/note= "N-myristoylated"
FT	Modified-site	230..343	/note= "N-myristoylated"
FT	Modified-site	105..126	/note= "Amidation"
FT	Peptide		/note= "predicted leucine zipper"
FN	WO200173051-A2.		
PD	04-OCT-2001.		
XX			
XX	26-MAR-2001; 2001WO-US09633.		
PR	24-MAR-2000; 2000US-192092P.		
XX			
PA	(MILL-) MILLENNIUM PHARM INC.		
XX			
PI	Meyers RA, Rudolph-Owen L, Macbeth KJ, Tsai F, Williamson M.		
DR	WPI: 2001-626267/72.		
XX			
DR	N-PSDB; AAH27094.		
XX			
PT	New polynucleotide encodes a polypeptide for the diagnosis and		
PT	treatment of cancer and to alter cellular lipid metabolism comprises a		
PT	polynucleotide encoding human acyltransferases designated 46743 and		
PT	27417		
XX			
PS	Claim 9; Fig 8A-B; 150pp; English.		
XX			
CC	The present sequence is that of a novel human protein, termed		
CC	27417, which shows the structural characteristics of members of the		
CC	acyltransferase family, including an acyltransferase domain. The		
CC	invention provides novel acyltransferase 27417 nucleic acids and		
CC	polypeptides, as well as methods for detecting their presence, and		
CC	methods for screening for compounds that modulate their expression		
CC	or activity. Such compounds can be used to treat conditions		
CC	related to aberrant activity or expression of 27417 protein or		
CC	nucleic acid, such as conditions involving deficient cellular		
CC	proliferation, migration and/or differentiation. Also provided are		
CC	methods for inhibiting the proliferation or migration, or inducing		
CC	the killing, of a 27417-expressing cell, e.g. a hyperproliferative		
CC	and/or metastatic cell, by contact with a compound that modulates		
CC	the activity or expression of the 27417 protein or nucleic acid.		
CC	The 27417-expressing cell is found in the healthy or diseased		
CC	heart, blood vessels, kidney, skeletal muscle, brain or liver, or		
CC	especially in a solid tumour, a soft tissue tumour or a metastatic		
CC	lesion, a sarcoma, a carcinoma or an adenocarcinoma, and in		
CC	particular in a hyperproliferative and/or metastatic cell found in		
CC	ovarian, brain, colon or lung cancer. The compounds may also be		
CC	used to modulate lipid metabolism in a 27417-expressing cell.		
CC	Methods of disease diagnosis, e.g. by determining the		
CC	presence of a genetic alteration in a 27417 polypeptide, and for		
CC	evaluating the efficacy of a treatment of a disorder, are also		
CC	provided.		
XX			
XX	Sequence 364 AA;		
0Y	Query Match	100.0%; Score 1871; DB 22; Length 364;	
	Best Local Similarity	100.0%; Pred. No. 2.le-189;	
	Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
7D	1 MRYLLPSVLLGTATYVLAWGWRLLSAPLPARYOALDRLVCVYSMTVEFFENTG 60		
	12 mryllpsvlllgtaptyvlaawgwrllsalfparfygaldldtlycvysmvtffentg 71		

QY	61	VOLLVLLGDLPEKRNKNTIITYLANHOSVDMIVADLIATRONALGHVRVYLKEGUKMLPIVGC	120
QY	72	VQIILLYGDLIPKXKNNIILYLANHGSYGVWIVADIIATIRQAIGHVTVLKEGILKWPILYGC	131
QY	121	YFAOHGGLYVRSKAFNFEKEMRNKLSQSYVAGCPMWLVLFPEGTRRNPQGTIVLSASQAF	180
DB	132	YFAQDQGLYVRSKFKFEKEMRNKLSGYVDAQPMYLVLFPEGTRYNPQGTIVLSAQSAF	191
QY	181	AAQGLAVLKVHVLTPRIKATHVAFDCKNLTDAIIVDTVYVYEGSKDDGGQRSPMTPEFL	240
DB	192	AAGTGLAVLKVHVLTPRIKATHVAFDCKNYLDIIVDTVYVYEGSKDDGGQRSPMTPEFL	251
QY	241	CKECKKHIIHIDRIDKKDVPDEQEHMRWYHEHPEELKDKMLIEFYSPSPERRKRPFGKS	300
DB	252	CKECKKHIIHIDRIDKKDVPDEQEHMRWYHEHPEELKDKMLIEFYSPSPERRKRPFGKS	311
QY	301	VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLYVTWTVYGTLLGLCLWYTIKA	353
DB	312	VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLYVTWTVYGTLLGLCLWYTIKA	364
RESULT	5		
AAU00665			
ID	AAU00665	standard; Protein; 364 AA.	
XX	AAU00665:		
AC			
XX	07-SEP-2001	(first entry)	
DE		Human lysophosphatidic acid acyltransferase isoform LPAAT-epsilon.	
XX		Lysophosphatidic acid acyltransferase epsilon; LPAAT-epsilon; LPA; PA;	
KW		lysophosphatidic acid; phosphatidic acid; acylation; cellular activation;	
KM		phospholipid signalling; mitogenesis; inflammation; autoimmune disease;	
XX		oncology; cancer; obesity; gene therapy.	
KX			
XX	Homo sapiens.		
OS			
XX	MO200134782-A1.		
PN			
XX	17-MAY-2001.		
PD			
XX	02-NOV-2000; 2000WO-US03193.		
PE			
XX	09-NOV-1999; 99US-0436919.		
PR			
XX	(CELL-) CELL THERAPEUTICS INC.		
PA			
XX	Leung DW;		
PL			
XX	WPI: 2001-335920/35.		
DR	N-PSDB; AAS00649.		
XX			
PT	Novel isolated human isoform of lysophosphatidic acid		
PT	acyltransferase-epsilon useful for diagnostic, therapeutic and		
PT	screening purposes -		
XX			
PS	Claim 1; Fig 1; 48pp; English.		
XX			
CC	The sequence represents a human lysophosphatidic acid acyltransferase		
CC	(LPAAT) isoform, LPAAT-epsilon. LPAAT catalyses the acylation of		
CC	lysophosphatidic acid (LPA) to phosphatidic acid (PA). LPA and PA have		
CC	been identified as phospholipid signalling molecules that affect a wide		
CC	range of biological responses. PA is involved in cellular activation and		
CC	mitogenesis. Compounds that block PA generation and hence diminish lipid		
CC	biosynthesis and the signal involved in cell activation are of		
CC	therapeutic interest in the areas of inflammation and oncology (e.g.		
CC	autoimmune diseases and cancer) as well as obesity treatment.		
CC	LPAAT-epsilon and its corresponding DNA can be used in screening assays		
CC	to detect agents that stimulate or inhibit the activity of LPAAT and,		
CC	therefore, PA. The DNA is useful in tests to detect the presence or		
CC	expression of LPAAT-epsilon in relation to certain diseases and		

CC conditions, and in disease prevention and treatment. The sequences of the  
CC invention are also useful for diagnosis of diseases and conditions in  
CC which the expression of LPAAT enzyme is abnormal.  
XX

Sequence 364 AA:

Query Match 100.0%; Score 1871; DB 22; Length 364;  
Best Local Similarity 100.0%; Pred. No. 2.1e-189;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTAPTVYLAAGVWRLLSAFLPARFYQALDDRLCYQSWVLEFFENYTG 60  
DB 12 MYLLPSVLLGTAPTVYLAAGVWRLLSAFLPARFYQALDDRLCYQSWVLEFFENYTG 71  
QY 61 VOILLGDLPRKKNENTIIYLANHSTVDWIVADILAIKRONALGHVRYVLEKGLKWLPLXGC 120  
DB 72 VQILLYGDLPKKNENTIIYLANHSTVDWIVADILAIKRONALGHVRYVLEKGLKWLPLXGC 131  
QY 121 YFAOHGIGYVKSRAKFNEMRNKLSQSYDAGTGMVLYFPECTRRYNPEQTWLSASQAF 180  
DB 132 YFAGHGGIYVKSRAKFNEMRNKLSQSYDAGTGMVLYFPECTRRYNPEQTWLSASQAF 191  
QY 181 AAQRGLAVLKHLTPRIKATHVAFDCMKNYLDATYDVTVVYEGKDDGGRRESPTWTEFL 240  
DB 192 AAGRGLAVLKHLTPRIKATHVAFDCMKNYLDATYDVTVVYEGKDDGGRRESPTWTEFL 251  
QY 241 CKECPRIHIDRIKDDVPEDEHMRRLHEFEIKDKMLIEFYSPDERRRKRPFGKS 300  
DB 252 CKECPRIHIDRIKDDVPEDEHMRRLHEFEIKDKMLIEFYSPDERRRKRPFGKS 311  
QY 301 VNSKISIKKTLPSMLISGLTAGMLMTDAGRKLVTWTVYGTLLGLWTYTKA 353  
DB 312 VNSKISIKKTLPSMLISGLTAGMLMTDAGRKLVTWTVYGTLLGLWTYTKA 364

RESULT 6

AAB93595  
ID AAB93595 standard; Protein; 353 AA.

AC AAB93595;

DT 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:13028.

KW Human; primer: detection; diagnosis; antisense therapy; gene therapy.

OS Homo sapiens.

XX EPI074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

XX 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

PT Primer sets for synthesizing polynucleotides, particularly the 5602  
PT full-length cDNAs defined in the specification, and for the detection  
PT and/or diagnosis of the abnormality of the proteins encoded by the  
PT full-length cDNAs -

XX Claim 8; SEQ ID 13028; 2537pp + CD ROW; English.

CC The present invention describes primer sets for synthesizing 5602  
CC full-length cDNAs defined in the specification. Where a primer set  
CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary  
CC to the complementary strand of a polynucleotide which comprises one of  
CC the 5602 nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in  
CC the specification. The primer sets can be used in antisense therapy and  
CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
CC represent oligonucleotides, all of which are used in the exemplification  
CC of the present invention.

Sequence 353 AA:

Query Match 99.8%; Score 1868; DB 22; Length 353;  
Best Local Similarity 99.7%; Pred. No. 4.2e-189;  
Matches 352; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTAPTVYLAAGVWRLLSAFLPARFYQALDDRLCYQSWVLEFFENYTG 60  
DB 1 MYLLPSVLLGTAPTVYLAAGVWRLLSAFLPARFYQALDDRLCYQSWVLEFFENYTG 60  
QY 61 VOILLGDLPRKKNENTIIYLANHSTVDWIVADILAIKRONALGHVRYVLEKGLKWLPLXGC 120  
DB 61 VQILLYGDLPKKNENTIIYLANHSTVDWIVADILAIKRONALGHVRYVLEKGLKWLPLXGC 120  
QY 121 YFAOHGIGYVKSRAKFNEMRNKLSQSYDAGTGMVLYFPECTRRYNPEQTWLSASQAF 180  
DB 121 YFAGHGGIYVKSRAKFNEMRNKLSQSYDAGTGMVLYFPECTRRYNPEQTWLSASQAF 180  
QY 181 AAQRGLAVLKHLTPRIKATHVAFDCMKNYLDATYDVTVVYEGKDDGGRRESPTWTEFL 240  
DB 181 AAGRGLAVLKHLTPRIKATHVAFDCMKNYLDATYDVTVVYEGKDDGGRRESPTWTEFL 240  
QY 241 CKECPRIHIDRIKDDVPEDEHMRRLHEFEIKDKMLIEFYSPDERRRKRPFGKS 300  
DB 241 CKECPRIHIDRIKDDVPEDEHMRRLHEFEIKDKMLIEFYSPDERRRKRPFGKS 300  
QY 301 VNSKISIKKTLPSMLISGLTAGMLMTDAGRKLVTWTVYGTLLGLWTYTKA 353  
DB 301 VNSKISIKKTLPSMLISGLTAGMLMTDAGRKLVTWTVYGTLLGLWTYTKA 353

RESULT 7

AA36728  
ID AA36728 standard; Protein; 352 AA.

AC AA36728;

DT 27-SEP-1999 (first entry)

DE Human Pgl protein sequence.

XX Pgl gene; diallelic marker; human; PSA; Pgl-related diallelic marker;  
XX cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.

XX Homo sapiens.

XX

PN WO932644-A2.  
XX  
PD 01-JUL-1999.  
XX  
PF 22-DEC-1998; 98WO-IB02133.  
XX  
PR 09-SEP-1998; 98US-0099658.  
PR 22-DEC-1997; 97US-0996306.  
XX  
PA (GEST ) GENSET.  
XX  
PI Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;  
XX WPI: 1999-405178/34.  
DR N-PSDB; AAZ00872.  
XX  
PT Use of a prostate cancer associated gene and biallelic markers  
XX derived from it  
PS  
XX Claim 7; Page 190; 385pp; English.  
XX  
XX The invention relates to a mammalian PGI gene and protein, and a set of  
CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
CC used in a hybridisation assay, a sequencing assay, or in an  
CC allele-specific amplification assay for determining the identity of a  
CC nucleotide at a PGI-related biallelic marker. The methods can be used to  
CC detect and to assess the risk of developing cancer or prostate cancer.  
CC Early-stage diagnosis of prostate cancer relies on prostate specific  
CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
CC to its inability to discriminate between malignant and non-malignant  
CC affections of the organ. A need exists for both a reliable diagnostic  
CC procedure which would enable early-stage diagnosis, and for preventative  
CC and curative treatments of the disease. The PGI gene can be used for  
CC detection of prostate cancer, and the risk of developing it in the  
CC future, and can also be used to determine therapies for the disease.  
XX  
SQ Sequence 352 AA:

Query Match 99.8%; Score 1867; DB 20; Length 352;  
Best Local Similarity 100.0%; Pred. No. 5.4e-189;  
Matches 352; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRYLLPSVLLGTAPTVVLAMGWRLSAFLPARFYQALDRLCYQYSWVLEFFENYTG 60  
DB 1 mryllpsvllgtaptyvvlamgwrrlsafllparfyqalddrlcyvqyswvleffenytg 60  
OY 61 VQILLYGDLPKKNENTITLANHGSTVDWIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120  
DB 61 vqilllygdldpkknentitlanhgstvwdvdlairgnalghvryvlekgikwlpilygc 120  
OY 121 YFAOHGIIYVRSARFNEKEEMRNKLOSVDAGTPMYLVITPEGRYRYPEDOTKIVSASQAF 180  
DB 121 yfahogiiyvrsarfnekeemrnklosyvdagtpmylvitfpegrtryrpeotkivsasaqaf 180  
OY 181 AAOGLAVLKLHVLPRIKATHVAFDCMKNYLDALIDYTVVYEGKDDGQORRESPTMEFL 240  
DB 181 aagrglavllkhlvprlkathvafdcmknyldaldydtvvveygddgqorrespmtmefl 240  
OY 241 CKECPKHHITDRDKDQVBEBOEHMRRWLEHREFTDKMLIEFYESPDPERRKRRFGKS 300  
DB 241 ckecpkhhitdrdkdqvbeboehmrrwlehereftdkmliefyespdperrkrfrpgks 300  
OY 301 VNSKLSIKTKTLPMSLISGLTAGMLMTDAGRKLYVNTWITGTLIGCLAWTIK 352  
DB 301 vnsklstiktklpsmlisgltagmlmtdagrklyvntwitygtlilgclawtlk 352

RESULT 8  
ID AAM39591 standard; protein; 353 AA.  
XX  
AC AAM39591;

XX 22-OCT-2001 (first entry)  
XX  
XX Human polypeptide SEQ ID NO 2736.  
DE  
XX  
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200153312-A1.  
PN  
XX  
XX 26-JUL-2001.  
PD  
XX  
XX 26-DEC-2000; 2000WO-US34263.  
PF  
XX  
XX 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-0552317.  
PR 09-JUL-2000; 2000US-0598042.  
PR 19-JUL-2000; 2000US-0620312.  
PR 03-AUG-2000; 2000US-0653450.  
PR 14-SEP-2000; 2000US-0662191.  
PR 19-OCT-2000; 2000US-0693036.  
PR 29-NOV-2000; 2000US-0727344.  
XX  
XX (HYSE-) HYSEQ INC.  
PA  
XX  
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX  
XX WPI: 2001-442253/47.  
DR N-PSDB; AAI58747.  
XX  
XX Novel nucleic acids and polypeptides, useful for treating disorders  
PT such as central nervous system injuries -  
PT  
XX  
PS Example 4: SEQ ID NO 2736; 10078bp; English.  
XX  
XX The invention relates to human nucleic acids (AA157798-AA161369) and  
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
CC immunosuppressant and cyostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression.  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.  
XX  
SQ Sequence 353 AA:

Query Match 99.5%; Score 1861; DB 22; Length 353;  
Best Local Similarity 99.2%; Pred. No. 2.3e-188;  
Matches 350; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRYLLPSVLLGTAPTVVLAMGWRLSAFLPARFYQALDRLCYQYSWVLEFFENYTG 60  
DB 1 mryllpsvllgtaptyvvlamgwrrlsafllparfyqalddrlcyvqyswvleffenytg 60  
OY 61 VQILLYGDLPKKNENTITLANHGSTVDWIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120  
DB 61 vqilllygdldpkknentitlanhgstvwdvdlairgnalghvryvlekgikwlpilygc 120

Db 61 vqlllygdipkknkennylianqstvdwivadilaalrqnalghvryvklwiplygc 120  
 QY 121 YFAOHGIGYVKSRAKFNKEMRNKLOSVDAGTPMVLVFPFGTRNPNPOTKVLASQAF 180  
 Db 121 yfaohgiglyvksrskfnekmrnklsqsyvdagtpmvlvfpfgtrypnpektvlssagaf 180  
 QY 181 AAORGLAVLKHVLTPIRIKATHVAFDCMKNYLDIAIVDTVYVYEGKDDGGRRESPTWTEFL 240  
 Db 181 aagrglavlkhvltprikathvafdomknyldaiydvtrvyegkddggrresptwtefl 240  
 QY 241 CKECKRIHHDRIKDDVPEOEHMRRWLHERFEIKDKMLIEFYSPDERRRKRPFGS 300  
 Db 241 ckecpkihhdridkddvpeegehmrrwlherfeikdkmliefesppdkrrkrfpgs 300  
 QY 301 VNSKLSIKKTLPSMLILSGITAGMLMTDAGRKLVTWYTWYTGTLGCLWYTIKA 353  
 Db 301 vnskslisktlpsmlllsgitagmlmtdagrklyvntwlytgcllwltika 353

## RESULT 9

AAM41377  
 ID AAM41377 standard; Protein: 372 AA.

AC AAM41377;  
 DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 6308.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia.

OS Homo sapiens.

PN WO200153312-A1.

PD 26-JUL-2001.

PF 26-DEC-2000; 2000WO-US34263.

PR 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.

PR 29-NOV-2000; 2000US-0727344.

DR WPI: 2001-442253/47.

DR N-PSDB; AA160533.

PT Novel nucleic acids and polypeptides, useful for treating disorders

PT such as central nervous system injuries -

XX Example 2; SEQ ID NO 6308; 10078pp; English.

CC The invention relates to human nucleic acids (AA157798-AA161369) and  
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and

CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokine activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemias and  
 CC C.N.S disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.

XX Sequence 372 AA:

Query Match 99.4%; Score 1860; DB 22: Length 372;  
 Best Local Similarity 99.7%; Pred. No. 3, 2e-188;  
 Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRYLPSVVLGTAIPYVLAAGVWRLLSAFLPARFYQALDDRLCYQSMVLEFFENYTG 60  
 Db 20 mryllpsvvligtapryvlaagvwrlsalfparfyqalddrlcyqsgmvleffenytg 79  
 QY 61 VOILLYGDLPKKNKENTIIYANHOSTVDWIVADILAIRQNALGHVRYVLEKGLWPLYGC 120  
 Db 80 vqlllygdipkknkennylianqstvdwivadilaalrqnalghvryvklwiplygw 139  
 QY 121 YFAOHGIGYVKSRAKFNKEMRNKLOSVDAGTPMVLVFPFGTRNPNPOTKVLASQAF 180  
 Db 140 yfaohgiglyvksrskfnekmrnklsqsyvdagtpmvlvfpfgtrypnpektvlssagaf 199  
 QY 181 AAORGLAVLKHVLTPIRIKATHVAFDCMKNYLDIAIVDTVYVYEGKDDGGRRESPTWTEFL 240  
 Db 200 aagrglavlkhvltprikathvafdomknyldaiydvtrvyegkddggrresptwtefl 259  
 QY 241 CKECKRIHHDRIKDDVPEOEHMRRWLHERFEIKDKMLIEFYSPDERRRKRPFGS 300  
 Db 260 ckecpkihhdridkddvpeegehmrrwlherfeikdkmliefesppdkrrkrfpgs 319  
 QY 301 VNSKLSIKKTLPSMLILSGITAGMLMTDAGRKLVTWYTWYTGTLGCLWYTIKA 353  
 Db 320 vnskslisktlpsmlllsgitagmlmtdagrklyvntwlytgcllwltika 372

## RESULT 10

AAY36751  
 ID AAY36751 standard; Protein: 315 AA.

AC AAY36751;

DT 27-SEP-1999 (first entry)

DE PGI splice variant allele protein sequence.

XX PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;

KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.

OS Homo sapiens.

PN WO9932644-A2.

PD 01-JUL-1999.

PF 22-DEC-1998; 98WO-1B02133.

PR 09-SEP-1998; 98US-0099658.

PR 22-DEC-1997; 97US-0996306.

PA (GSEST ) GENSEST.

PI Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;

DR WPI: 1999-405178/34.

DR N-PSDB; AA200977.

XX Use of a prostate cancer associated gene and biallelic markers  
PT derived from it  
XX  
XX  
PS Claim 7: Page 253-254; 385pp: English.  
XX  
CC The invention relates to a mammalian PGI gene and protein, and a set of  
CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
CC used in a hybridisation assay, a sequencing assay, or in an  
CC allele-specific amplification assay for determining the identity of a  
CC nucleotide at a PGI-related biallelic marker. The methods can be used to  
CC detect and to assess the risk of developing cancer or prostate cancer.  
CC Early-stage diagnosis of prostate cancer relies on prostate specific  
CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
CC to its inability to discriminate between malignant and non-malignant  
CC affections of the organ. A need exists for both a reliable diagnostic  
CC procedure which would enable early-stage diagnosis, and for preventative  
CC and curative treatments of the disease. The PGI gene can be used for  
CC detection of prostate cancer, and the risk of developing it in the  
CC future, and can also be used to determine therapies for the disease.  
XX  
SQ Sequence 315 AA:  
  
Query Match 87.9%; Score 1644; DB 20; Length 315;  
Best Local Similarity 89.2%; Pred. No. 2e-165;  
Matches 315; Conservative 0; Mismatches 0; Indels 38; Gaps 1;  
  
QY 1 MRLTSLSVLLGTAPTYVLAAGWRLSALPRFYQALDDRLCYQSWLFEFFENYTG 60  
DB 1 MRLTSLSVLLGTAPTYVLAAGWRLSALPRFYQALDDRLCYQSWLFEFFENYTG 60  
QY 61 VOILLYGDLPKKNENIYYLANHOSTVDWIYADILAIROMALGHVRYLKEGLKWLPLYGC 120  
DB 61 VQILLYGDLPKKNENIYYLANHGSTVDWIYADILAIROMALGHVRYLKEGLKWLPLYGC 120  
QY 121 YFAOHGSIYKRSKAKFNEMKRNKLOSVDAGTPMTLVIFPECTRYNPEQTKVLSASQAF 180  
DB 121 YFAOHGSIYKRSKAKFNEMKRNKLOSVDAGTPMTLVIFPECTRYNPEQTKVLSASQAF 180  
QY 121 YFAOHGSIYKRSKAKFNEMKRNKLOSVDAGTPMTLVIFPECTRYNPEQTKVLSASQAF 180  
DB 121 YFAOHGSIYKRSKAKFNEMKRNKLOSVDAGTPMTLVIFPECTRYNPEQTKVLSASQAF 180  
QY 181 AAOAGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTYVYEGKDDGCGORRESPTMEFL 240  
DB 181 AAOAGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTYVYEGKDDGCGORRESPTMEFL 240  
QY 241 CKCEPKIHHIDRIKDDVPEQEHMRKWLHERFEIKDKMLIEFYESPDPERRKRPFGKS 300  
DB 241 CKCEPKIHHIDRIKDDVPEQEHMRKWLHERFEIKDKMLIEFYESPDPERRKRPFGKS 300  
QY 301 VNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLYVNTWYIGTLGCLWTTIKA 353  
DB 301 VNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLYVNTWYIGTLGCLWTTIKA 353  
DB 263 VNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLYVNTWYIGTLGCLWTTIKA 315  
  
RESULT 11  
AA36741  
ID AAY36741 standard; Protein: 354 AA.  
XX  
AC AAY36741;  
XX  
DT 27-SEP-1999 (first entry)  
XX  
DE Mouse PGI protein sequence.  
XX  
KW PGI gene: biallelic marker; mouse; PSA; PGI-related biallelic marker;  
KM cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.  
XX  
OS Mus musculus.  
XX  
PN WO9332644-A2.  
XX  
PD 01-JUL-1999.  
XX  
PF 22-DEC-1998; 98WO-IB02133.

XX 09-SEP-1998; 98US-0099658.  
PR 22-DEC-1997; 97US-0996306.  
XX  
XX  
PA (GEST ) GENSET.  
XX  
PI Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;  
XX WPI: 1999-405178/34.  
DR N-PSDB: AAZ00928, AAZ01027.  
XX  
PT Use of a prostate cancer associated gene and biallelic markers  
PT derived from it  
XX  
PS Claim 7: Page 215; 385pp: English.  
XX  
CC The invention relates to a mammalian PGI gene and protein, and a set of  
CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
CC used in a hybridisation assay, a sequencing assay, or in an  
CC allele-specific amplification assay for determining the identity of a  
CC nucleotide at a PGI-related biallelic marker. The methods can be used to  
CC detect and to assess the risk of developing cancer or prostate cancer.  
CC Early-stage diagnosis of prostate cancer relies on prostate specific  
CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
CC to its inability to discriminate between malignant and non-malignant  
CC affections of the organ. A need exists for both a reliable diagnostic  
CC procedure which would enable early-stage diagnosis, and for preventative  
CC and curative treatments of the disease. The PGI gene can be used for  
CC detection of prostate cancer, and the risk of developing it in the  
CC future, and can also be used to determine therapies for the disease.  
XX  
SQ Sequence 354 AA:  
  
Query Match 83.8%; Score 1568.5; DB 20; Length 354;  
Best Local Similarity 81.4%; Pred. No. 2.3e-157;  
Matches 288; Conservative 39; Mismatches 26; Indels 1; Gaps 1;  
  
QY 1 MRLTSLSVLLGTAPTYVLAAGWRLSALPRFYQALDDRLCYQSWLFEFFENYTG 60  
DB 1 MRLTSLSVLLGTAPTYVLAAGWRLSALPRFYQALDDRLCYQSWLFEFFENYTG 60  
QY 61 VOILLYGDLPKKNENIYYLANHOSTVDWIYADILAIROMALGHVRYLKEGLKWLPLYGC 120  
DB 61 VQILLYGDLPKKNENIYYLANHGSTVDWIYADILAIROMALGHVRYLKEGLKWLPLYGC 120  
QY 121 YFAOHGSIYKRSKAKFNEMKRNKLOSVDAGTPMTLVIFPECTRYNPEQTKVLSASQAF 180  
DB 121 YFAOHGSIYKRSKAKFNEMKRNKLOSVDAGTPMTLVIFPECTRYNPEQTKVLSASQAF 180  
QY 121 YFAOHGSIYKRSKAKFNEMKRNKLOSVDAGTPMTLVIFPECTRYNPEQTKVLSASQAF 180  
DB 121 YFAOHGSIYKRSKAKFNEMKRNKLOSVDAGTPMTLVIFPECTRYNPEQTKVLSASQAF 180  
QY 181 AAOAGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTYVYEGKDDGCGORRESPTMEFL 240  
DB 181 AAOAGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTYVYEGKDDGCGORRESPTMEFL 240  
QY 240 LCKCEPKIHHIDRIKDDVPEQEHMRKWLHERFEIKDKMLIEFYESPDPERRKRPFGK 299  
DB 240 LCKCEPKIHHIDRIKDDVPEQEHMRKWLHERFEIKDKMLIEFYESPDPERRKRPFGK 299  
QY 300 SVNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLYVNTWYIGTLGCLWTTIKA 353  
DB 301 SVNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLYVNTWYIGTLGCLWTTIKA 354  
  
RESULT 12  
AA36752  
ID AAY36752 standard; Protein: 300 AA.  
XX  
AC AAY36752;  
XX  
DT 27-SEP-1999 (first entry)  
XX  
DE PGI splice variant allele protein sequence.  
XX



KW PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;  
KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.

XX Homo sapiens.

XX W09932644-A2.

XX 01-JUL-1999.

XX 22-DEC-1998; 98MO-IB02133.

XX 09-SEP-1998; 98US-0099658.

XX 22-DEC-1997; 97US-0996306.

XX (GEST ) GENSEP.

XX Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;

XX WPI; 1999-405178/34.

XX N-PSDB; AAZ00978.

XX Use of a prostate cancer associated gene and biallelic markers  
XX derived from it

XX Claim 7; Page 254-255; 385pp; English.

XX The invention relates to a mammalian PGI gene and protein, and a set of  
XX PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
XX used in a hybridisation assay, a sequencing assay, or in an  
XX allele-specific amplification assay for determining the identity of a  
XX nucleotide at a PGI-related biallelic marker. The methods can be used to  
XX detect and to assess the risk of developing cancer or prostate cancer.  
XX Early-stage diagnosis of prostate cancer relies on prostate specific  
XX antigen (PSA) dosage. However, the effectiveness of this is limited due  
XX to its inability to discriminate between malignant and non-malignant  
XX affections of the organ. A need exists for both a reliable diagnostic  
XX procedure which would enable early-stage diagnosis, and for preventative  
XX and curative treatments of the disease. The PGI gene can be used for  
XX detection of prostate cancer, and the risk of developing it in the  
XX future, and can also be used to determine therapies for the disease.

XX Sequence 300 AA;

Query Match 83.1%; Score 1555.5; DB 20; Length 300;

Best Local Similarity 85.0%; Pred. No. 4.3e-156; Indels 53; Gaps 1;

Matches 300; Conservative 0; Mismatches 0; Indels 53; Gaps 1;

QY 1 MRYLLPSVVLGTAPTYVLAMGVMRLLSAFLPARYQALDRLCYVOSMVLFFFEYNTG 60

DB 1 mryllpsvvlgtaptlylavgvwrlsalfparfygalddrlcyvgsmviffenytg 60

QY 61 VOILLYGLPKKKNIIYLANHOSVVDWIVADILAIRONALGHVRYLKEGKMLPLYGC 120

DB 61 vqillyglpkkeniilylanhqstvdwlvadilairqalghvrylkegkmlplygc 120

QY 121 YFAOHGIGYVRSKAFNEKEMNKLOSVDAGTPMYLVIFPEGSTRYNPEOTVLSASQAF 180

DB 121 yfaogigiyvrsakfnekenmklosyvdagtpmylvifpegstrynpeotvlsasqaf 180

QY 181 AAORGLAVLKHVLPRIKATHVAFDCMKNYLDAIVDVTVVYEGKDDGGRRSPMTTEFL 240

DB 181 aagr-----efl 187

QY 241 CKECPKIHIDRIDKKVPEEOEHMRRLHERFEIKDKMLIEFYESPDPERRKRFPKGS 300

DB 188 ckecpkihhdridkkvpeegehmrrwlherfeikdkmliefespderrkrfpgks 247

QY 301 VNSKSIKIKTLPSMLILSGLTAGMLATDAGKRLVNTWITYGTLGCLWWTYTA 353

DB 248 vnsksikiktlpsmlilsgltagmlatdagkrlvntwitygtlglclwtlyka 300

RESULT 13  
AAV36744  
ID AAV36744 standard; Protein; 291 AA.

XX AAY36744;

XX 27-SEP-1999 (first entry)

XX PGI splice variant allele protein sequence.

XX PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;

XX cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.

XX Homo sapiens.

XX W09932644-A2.

XX 01-JUL-1999.

XX 22-DEC-1998; 98MO-IB02133.

XX 09-SEP-1998; 98US-0099658.

XX 22-DEC-1997; 97US-0996306.

XX (GEST ) GENSEP.

XX Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;

XX WPI; 1999-405178/34.

XX N-PSDB; AAZ00970.

XX Use of a prostate cancer associated gene and biallelic markers  
XX derived from it

XX Claim 7; Page 250-251; 385pp; English.

XX The invention relates to a mammalian PGI gene and protein, and a set of  
XX PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
XX used in a hybridisation assay, a sequencing assay, or in an  
XX allele-specific amplification assay for determining the identity of a  
XX nucleotide at a PGI-related biallelic marker. The methods can be used to  
XX detect and to assess the risk of developing cancer or prostate cancer.  
XX Early-stage diagnosis of prostate cancer relies on prostate specific  
XX antigen (PSA) dosage. However, the effectiveness of this is limited due  
XX to its inability to discriminate between malignant and non-malignant  
XX affections of the organ. A need exists for both a reliable diagnostic  
XX procedure which would enable early-stage diagnosis, and for preventative  
XX and curative treatments of the disease. The PGI gene can be used for  
XX detection of prostate cancer, and the risk of developing it in the  
XX future, and can also be used to determine therapies for the disease.

XX Sequence 291 AA;

Query Match 80.1%; Score 1498; DB 20; Length 291;

Best Local Similarity 82.4%; Pred. No. 5.1e-150; Indels 62; Gaps 1;

Matches 291; Conservative 0; Mismatches 0; Indels 62; Gaps 1;

QY 1 MRYLLPSVVLGTAPTYVLAMGVMRLLSAFLPARYQALDRLCYVOSMVLFFFEYNTG 60

DB 1 mryllpsvvlgtaptlylavgvwrlsalfparfygalddrlcyvgsmviffenytg 60

QY 61 VOILLYGLPKKKNIIYLANHOSVVDWIVADILAIRONALGHVRYLKEGKMLPLYGC 120

DB 61 v----- 61

QY 121 YFAOHGIGYVRSKAFNEKEMNKLOSVDAGTPMYLVIFPEGSTRYNPEOTVLSASQAF 180

DB 62 ---qnggilyvrsakfnekenmklosyvdagtpmylvifpegstrynpeotvlsasqaf 118

QY 181 AAORGLAVLKHVLPRIKATHVAFDCMKNYLDAIVDVTVVYEGKDDGGRRSPMTTEFL 240

DB 119 aagrglavlkhvlprikathvafdcmknyldaidvvtvvyegkddggrspmttefl 178

Qy	241	CKECPKIHIDRIDKDDVEEDGHRMRMLHEFEITKDKMLIEFESDPERRKRPPEKS	300
Db	179	cKECPKIHhIdrIdKkdvpeEGemrIwIherfeIkdkmlIefesDPerrkrIpgs	238
Qy	301	VNSKSIKTKTLPsMLISGLTAGKLMWDGGRKLYNTWTWIGTLIGLWTIKA	353
Db	239	vnsKslsktlcIpsmLIISgltaGmImtdagrkIyVntwIygtLIgclwIcIka	291
RESULT 14			
AAE09653			
ID	AAE09653 standard; Protein; 269 AA.		
XX			
AC	AAE09653;		
XX			
DT	19-NOV-2001 (first entry)		
XX			
DE	Human gene 15 encoded lipid metabolism protein HHASB32, SEQ ID NO:47.		
KW	Human; lipid metabolism protein; noctropic; neuroprotective; cardiant;		
KW	cerebroprotective; hepatotropic; antidiabetic; ophthalmic; nephrotropic;		
KW	immune disorder; autoimmune disease; rheumatoid arthritis; glossitis;		
KW	systemic lupus erythematosus; conjunctivitis; inflammatory disorder;		
KW	respiratory disorder; asthma; allergy; CMS disorder; Alzheimer's disease;		
KW	Parkinson's disease; atherosclerosis; cardiovascular disorder; cancer;		
KW	coronary disease; familial hypercholesterolemia; hyperlipidaemia;		
KW	haematopoietic disorder; hypolipidaemia; lipidosi; Gaucher's disease;		
KW	Tay-Sach's disease; mental retardation; gene therapy; antisense therapy.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200155203-A1.		
XX			
PD	02-AUG-2001.		
XX			
PF	17-JAN-2001; 2001WO-US01327.		
XX			
PR	31-JAN-2000;	2000US-0179065.	
PR	04-FEB-2000;	2000US-0180628.	
PR	24-FEB-2000;	2000US-0184664.	
PR	02-MAR-2000;	2000US-0186350.	
PR	16-MAR-2000;	2000US-0198974.	
PR	17-MAR-2000;	2000US-0190076.	
PR	18-APR-2000;	2000US-0198123.	
PR	19-MAY-2000;	2000US-0205515.	
PR	07-JUN-2000;	2000US-0209467.	
PR	28-JUN-2000;	2000US-0214886.	
PR	30-JUN-2000;	2000US-0215135.	
PR	07-JUL-2000;	2000US-0216647.	
PR	07-JUL-2000;	2000US-0216680.	
PR	11-JUL-2000;	2000US-0217487.	
PR	11-JUL-2000;	2000US-0217496.	
PR	14-JUL-2000;	2000US-0218290.	
PR	26-JUL-2000;	2000US-0220963.	
PR	26-JUL-2000;	2000US-0220964.	
PR	14-AUG-2000;	2000US-0224518.	
PR	14-AUG-2000;	2000US-0224519.	
PR	14-AUG-2000;	2000US-0225213.	
PR	14-AUG-2000;	2000US-0225214.	
PR	14-AUG-2000;	2000US-0225266.	
PR	14-AUG-2000;	2000US-0225267.	
PR	14-AUG-2000;	2000US-0225268.	
PR	14-AUG-2000;	2000US-0225270.	
PR	14-AUG-2000;	2000US-0225447.	
PR	14-AUG-2000;	2000US-0225477.	
PR	14-AUG-2000;	2000US-0225757.	
PR	14-AUG-2000;	2000US-0225758.	
PR	14-AUG-2000;	2000US-0225759.	
PR	18-AUG-2000;	2000US-0226279.	
PR	22-AUG-2000;	2000US-0226681.	
PR	22-AUG-2000;	2000US-0226686.	
PR	22-AUG-2000;	2000US-0227182.	
PR	23-AUG-2000;	2000US-0227009.	

17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-02559678.  
 XX  
 PA (HDMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX  
 DR WPI: 2001-457716/49.  
 DR N-PSDB: AAD16620.  
 XX  
 PT Isolated lipid metabolism polypeptide for screening to identify  
 PT antagonists and agonists that may enhance or block activities mediated  
 PT by lipid metabolism proteins and also for testing and detection e.g.  
 PT diagnosis -  
 PS  
 PS Claim 11; SEQ ID NO 47; 453bp; English.  
 XX  
 XX AAD16606-AAD16627 represent cDNAs corresponding to human lipid metabolism  
 CC protein genes, and AA09639-AA09660 represent the proteins they encode.  
 CC AAD16628-AAD16647 represent human lipid metabolism genomic DNAs. The  
 CC genes and their corresponding proteins are useful for preventing,  
 CC treating or ameliorating medical conditions such as immune disorders,  
 CC inflammatory disorders (e.g., conjunctivitis, glossitis), autoimmune  
 CC diseases (e.g., rheumatoid arthritis, systemic lupus erythematosus),  
 CC respiratory disorders (asthma, allergy), CNS disorders (e.g., Alzheimer's  
 CC disease, Parkinson's disease), haematopoietic disorders, cardiovascular  
 CC disorders (e.g., atherosclerosis, coronary disease), hyperlipidaemia,  
 CC hypolipidaemia, lipidosis, familial hypercholesterolaemia, Gaucher's  
 CC disease, Tay-Sachs disease, diabetes mellitus, cancer and other enzyme  
 CC abnormalities. They are used for the prevention of mental retardation  
 CC and the enlargement of spleen and liver. They are used in gene therapy  
 CC and antisense therapy. The present sequence represents a human lipid  
 CC metabolism protein of the invention.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 XX Sequence 269 AA:

Query Match 76.3%; Score 1427; DB 22; Length 269;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-142;  
 Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 TWDWIVADILAIRONALGHVRYLVKGLKWLPLVGYCYFAOHGIGIYVRSARAKNEKEMRNK 144  
 DB 1 twdwivadilaironalghvrylvkglkwlplvgcyfahqhg9lykrsakfnekemrnk 60

QY 145 LQSYVDAGTPMYLVIFPEEGTRYNPEQTKVLSASQAFEAORGLAVLKHVLPRIKATNVAF 204  
 DB 61 lqsyvdagtpmylvifpeegtrynpeqtkvlsasqafaagqglavlkhvlprikatnvaf 120  
 QY 205 DCKKNYLDALDYVTYVEGKDDGQRRSEPTMEFLCKECPKIHIDRIDKKDVEEQE 264  
 DB 121 dckknlyldaldyvtvyegkddgqrrseptmeflckeckpihridridkkdveepe 180  
 QY 265 HMRRLWHERREIKDKMLIERESPDPERRKRFQGSVNSKLSIKTLPSMLISGLTAGM 324  
 DB 181 hmrwlwherreikdkmliefespdperrkrfpgksvnsklsiktlpsmlisgltagm 240  
 QY 325 LMTDAGRKLXVNTWITGTLGCLMTYTKA 353  
 DB 241 lmtdagrklyvntwlygtlglcylwtika 269

RESULT 15  
 AAY36745  
 ID AAY36745 standard; Protein; 261 AA.  
 XX  
 AC AAY36745;  
 XX  
 DT 27-SEP-1999 (first entry)  
 XX  
 DE PGI splice variant allele protein sequence.  
 XX  
 XX PGI gene; biallelic marker; human; PSA: PGI-related biallelic marker;  
 KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.  
 KM  
 XX Homo sapiens.  
 OS  
 XX  
 XX W09932644-A2.  
 PN  
 XX  
 PD 01-JUL-1999.  
 XX  
 PF 22-DEC-1998; 98WO-1B02133.  
 XX  
 PR 09-SEP-1998; 98US-0099658.  
 PR 22-DEC-1997; 97US-0996306.  
 XX  
 XX  
 PA (GEST ) GENSET.  
 XX  
 PI Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;  
 DR WPI: 1999-405178/34.  
 DR N-PSDB: AA200971.  
 XX  
 PT Use of a prostate cancer associated gene and biallelic markers  
 PT derived from it  
 PS  
 PS Claim 7; Page 251; 385bp; English.  
 XX  
 XX The invention relates to a mammalian PGI gene and protein, a set of  
 CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
 CC used in a hybridisation assay, a sequencing assay, or in an  
 CC allele-specific amplification assay for determining the identity of a  
 CC nucleotide at a PGI-related biallelic marker. The methods can be used to  
 CC detect and to assess the risk of developing cancer or prostate cancer.  
 CC Early-stage diagnosis of prostate cancer relies on prostate specific  
 CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
 CC to its inability to discriminate between malignant and non-malignant  
 CC affections of the organ. A need exists for both a reliable diagnostic  
 CC procedure which would enable early-stage diagnosis, and for preventative  
 CC and curative treatments of the disease. The PGI gene can be used for  
 CC detection of prostate cancer, and the risk of developing it in the  
 CC future, and can also be used to determine therapies for the disease.  
 XX  
 XX Sequence 261 AA:

Query Match 70.8%; Score 1325; DB 20; Length 261;

Best Local Similarity 73.9%; Pred. No. 9.4e-132;  
Matches 261; Conservative 0; Mismatches 0; Indels 92; Gaps 1;

```
QY      1 MRYLLPSVVLGTAPTYVLAWGVRLSAFLPARFYQALDDRLCYVQSMVLPFFENYTG 60
      |||
Db      1 mryllpsvvl|gtaptvylawgvrrllsaflparfyqalddrlcyvqsmv|lffenytg 60
QY      61 VOILLYGDLPNNKENIYYLANHGSTVDMIVADILAIRONALGHVRYVLKGLKMLPLYGC 120
      ||
Db      61 vq----- 62
QY      121 YFAGHGIIYVKRSKAFNEKEKMRNKLQSYVDAGTPMILYIFPESTRYNPQTKVLSASQAF 180
      |||
Db      63 -----mylvlfpegrlrynpqtkvl|saqaf 88
QY      181 AAORGLAVLKHYLPRIKATHVAFDCMKKNYLDATYDVTVVEGKDDGQRESPTMTEFL 240
      |||
Db      89 aagrglavlkhvl|prlkathvafdcmkkn|ldaldvltvyegkddgqrrtespuncetfl 148
QY      241 CKECPKIHIDRIDKKDVPEQEHMRRMLHEREIKDKMLIEFYESPDPERRRRFPFGKS 300
      |||
Db      149 ckecpk|hhidridkdkdvpeeqehmrrml|herfelkdkml|iefyespdperrrr|frfgks 208
QY      301 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLYVNTWYIYGTLLGCLMWTIKA 353
      |||
Db      209 vnsklslkclpsmll|sgltagmlmtdagrk|lyvntwlygtll|gclmwtika 261
```

Search completed: August 28, 2002, 11:07:33  
Job time: 477 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:03:21 ; Search time 31 Seconds  
(without alignments)  
278.137 Million cell updates/sec

Title: US-09-853-526-4

Perfect score: 1871  
Sequence: 1 MRYLPSVVLGTAPRYVLA.....YVNTWIXGTLGLWVTIXA 353

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 segs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA.\*

1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCITUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfills1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1871	100.0	353	2	US-08-996-306-4
2	1871	100.0	353	4	US-09-338-907-4
3	1871	100.0	353	4	US-09-218-207-4
4	1871	100.0	364	2	US-08-996-306-5
5	1871	100.0	364	4	US-09-338-907-5
6	1871	100.0	364	4	US-09-218-207-5
7	1644	87.9	315	4	US-09-338-907-134
8	1644	87.9	315	4	US-09-218-207-134
9	1568.5	83.8	354	4	US-09-338-907-74
10	1568.5	83.8	354	4	US-09-218-207-74
11	1555.5	83.1	300	4	US-09-338-907-135
12	1555.5	83.1	300	4	US-09-218-207-135
13	1498	80.1	291	4	US-09-338-907-127
14	1498	80.1	291	4	US-09-218-207-127
15	1325	70.8	261	4	US-09-338-907-128
16	1325	70.8	261	4	US-09-218-207-128
17	1203	64.3	228	4	US-09-338-907-70
18	1203	64.3	228	4	US-09-218-207-70
19	1182.5	63.2	238	4	US-09-338-907-126
20	1182.5	63.2	238	4	US-09-218-207-126
21	976	52.2	185	4	US-09-338-907-136
22	976	52.2	185	4	US-09-218-207-136
23	820	43.8	182	4	US-09-338-907-133
24	820	43.8	182	4	US-09-218-207-133
25	449	24.0	97	4	US-09-338-907-132
26	449	24.0	97	4	US-09-218-207-132
27	332	17.7	77	4	US-09-338-907-125

28	332	17.7	77	4	US-09-218-207-125	Sequence 125, App
29	330	17.6	66	4	US-09-338-907-131	Sequence 131, App
30	330	17.6	66	4	US-09-218-207-131	Sequence 131, App
31	329	17.6	68	4	US-09-338-907-130	Sequence 130, App
32	329	17.6	68	4	US-09-218-207-130	Sequence 130, App
33	329	17.6	90	4	US-09-338-907-129	Sequence 129, App
34	329	17.6	90	4	US-09-218-207-129	Sequence 129, App
35	328	17.5	374	2	US-08-454-267-2	Sequence 2, App1
36	328	17.5	374	2	US-08-941-319-2	Sequence 2, App1
37	328	17.5	374	2	US-08-941-319-2	Sequence 2, App1
38	328	17.5	374	2	US-08-941-319-6	Sequence 6, App1
39	328	17.5	374	4	US-09-035-098-2	Sequence 2, App1
40	328	17.5	374	4	US-09-035-098-6	Sequence 6, App1
41	328	17.5	374	4	US-09-215-252-5	Sequence 5, App1
42	268.5	14.4	378	4	US-09-215-252-17	Sequence 17, App1
43	240	12.8	376	4	US-09-215-252-13	Sequence 13, App1
44	239	12.8	295	2	US-08-454-267-7	Sequence 7, App1
45	239	12.8	295	2	US-08-941-319-7	Sequence 7, App1

#### ALIGNMENTS

RESULT 1  
US-08-996-306-4  
; Sequence 4, Application US/08996306  
; Patent No. 5945522  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Chumakov, Ilya  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Bougenelert, Lydie  
; TITLE OF INVENTION: Prostate cancer gene  
; NUMBER OF SEQUENCES: 68  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobb, Martens, Olson & Bear  
; STREET: 501 West Broadway  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92101-3505  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Win95  
; SOFTWARE: Word  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/996,306  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israelson, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: GENSET.018A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 353 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: SINGLE  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: PROTEIN  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: potential Transmembrane helix  
; LOCATION: 1..33  
; IDENTIFICATION METHOD: Rao and Argos method  
; FEATURE:  
; NAME/KEY: potential Transmembrane helix  
; LOCATION: 4..20

IDENTIFICATION METHOD: Klein, Kanehisa and Delisi method  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 4..24  
IDENTIFICATION METHOD: Eisenberg,Schwarz,Komarony  
IDENTIFICATION METHOD: and Wall method  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 12  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 50..70  
IDENTIFICATION METHOD: Eisenberg,Schwarz,Komarony and Wall method  
FEATURE:  
NAME/KEY: potential N-glycosylation site  
LOCATION: 57  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 76..96  
IDENTIFICATION METHOD: Eisenberg,Schwarz,Komarony and Wall method  
FEATURE:  
NAME/KEY: potential Tyrosine kinase phosphorylation site  
LOCATION: 78  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Casein kinase II phosphorylation site  
LOCATION: 84  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Leucine zipper pattern  
LOCATION: 94..115  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 119  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 133  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Casein kinase II phosphorylation site  
LOCATION: 147  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 194  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Tyrosine kinase phosphorylation site  
LOCATION: 215  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Tyrosine sulfatation site  
LOCATION: 221  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential CAMP- and cGMP-dependent protein kinase phosphorylation site  
LOCATION: 223  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Casein kinase II phosphorylation site  
LOCATION: 235  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 306  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 310..330

IDENTIFICATION METHOD: Eisenberg,Schwarz,Komarony and Wall method  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 319  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 323  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Amidation site  
LOCATION: 329  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 333..353  
IDENTIFICATION METHOD: Eisenberg,Schwarz,Komarony and Wall method  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 341  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 350  
IDENTIFICATION METHOD: prosite match  
US-08-996-306-4  
Query Match 100.0%; Score 1871; DB 2; Length 353;  
Best Local Similarity 100.0%; Pred. No. 1.9e-194;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRYLPSVVLGTAPTYVLAMGVRLISAFIPAFYQALDDRLCYQYOSNVLEFFENYTG 60  
DB 1 MRYLPSVVLGTAPTYVLAMGVRLISAFIPAFYQALDDRLCYQYOSNVLEFFENYTG 60  
QY 61 VOILLYGDLPKKNENIYYLANHOSYDVIADILAIRONALGHVRYVLEKGLMPLYGC 120  
DB 61 VOILLYGDLPKKNENIYYLANHOSYDVIADILAIRONALGHVRYVLEKGLMPLYGC 120  
QY 121 YFAOHGSIYVRSKAFNEKEMRNKLOSVDAGTPMVLVIFPESTRNPBOTKVLASQAF 180  
DB 121 YFAOHGSIYVRSKAFNEKEMRNKLOSVDAGTPMVLVIFPESTRNPBOTKVLASQAF 180  
QY 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYIDAIVDVTVVYESGDKDGGORRESPTMTEFL 240  
DB 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYIDAIVDVTVVYESGDKDGGORRESPTMTEFL 240  
QY 241 CKECPKIHIDRIDKKDVPEDOEHHMRMLHERFEIKDKMLIEFYSPDERKRRPGKS 300  
DB 241 CKECPKIHIDRIDKKDVPEDOEHHMRMLHERFEIKDKMLIEFYSPDERKRRPGKS 300  
QY 301 VNSKLSIKRTLPMLISGLTAGMLMTDAGRKYVNTWYIGTLGLMTYTKA 353  
DB 301 VNSKLSIKRTLPMLISGLTAGMLMTDAGRKYVNTWYIGTLGLMTYTKA 353  
RESULT 2  
US-09-338-907-4  
Sequence 4, Application US/09338907  
Patent No. 6265346  
GENERAL INFORMATION:  
APPLICANT: Cohen, Daniel  
APPLICANT: Blumenfeld, Marla  
APPLICANT: Ilya, Chumakov  
APPLICANT: Bougueleret, Lydie  
TITLE OF INVENTION: PROSTATE CANCER GENE  
FILE REFERENCE: GENSET.18CP1CP  
CURRENT APPLICATION NUMBER: US/09/338,907  
CURRENT FILING DATE: 1999-06-23  
EARLIER APPLICATION NUMBER: 08/996,306  
EARLIER FILING DATE: 1997-12-22  
EARLIER APPLICATION NUMBER: 60/099,658

EARLIER FILING DATE: 1998-09-09  
EARLIER APPLICATION NUMBER: 09/218,207  
EARLIER FILING DATE: 1998-12-22  
NUMBER OF SEQ ID NOS: 578  
SOFTWARE: Patent.pm  
SEQ ID NO 4  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 1..33  
OTHER INFORMATION: Rao and Argos identification method, potential helix  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 4..20  
OTHER INFORMATION: Klein, Kanehisa and Delisi identification method, potential  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 4..24  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F  
FEATURE:  
NAME/KEY: MYRISTATE  
LOCATION: 12..16  
OTHER INFORMATION: Prosite match  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 50..70  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F  
FEATURE:  
NAME/KEY: CARBOHYD  
LOCATION: 57..59  
OTHER INFORMATION: Prosite match  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 76..96  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 78  
OTHER INFORMATION: potential Tyrosine kinase site, Prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 84  
OTHER INFORMATION: potential caseine kinase II site, Prosite match  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 94..115  
OTHER INFORMATION: potential leucine zipper site, Prosite match  
FEATURE:  
NAME/KEY: MYRISTATE  
LOCATION: 119..123  
OTHER INFORMATION: potential site, Prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 133  
OTHER INFORMATION: potential protein kinase C, Prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 147  
OTHER INFORMATION: potential caseine kinase II site, Prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 194  
OTHER INFORMATION: potential protein kinase C, Prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 215  
OTHER INFORMATION: potential Tyrosine kinase site, Prosite match  
FEATURE:  
NAME/KEY: SULFATATION  
LOCATION: 221  
OTHER INFORMATION: Prosite match

FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 223  
OTHER INFORMATION: potential cAMP and cGMP dependant protein kinase site,  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 235  
OTHER INFORMATION: potential caseine kinase II site, Prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 306  
OTHER INFORMATION: potential protein kinase C, Prosite match  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 310..330  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
FEATURE:  
NAME/KEY: MYRISTATE  
LOCATION: 319..323  
OTHER INFORMATION: Prosite match  
FEATURE:  
NAME/KEY: MYRISTATE  
LOCATION: 323..327  
OTHER INFORMATION: Prosite match  
FEATURE:  
NAME/KEY: AMIDATION  
LOCATION: 329  
OTHER INFORMATION: Prosite match  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 333..353  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
FEATURE:  
NAME/KEY: MYRISTATE  
LOCATION: 341..345  
OTHER INFORMATION: Prosite match  
FEATURE:  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 350  
OTHER INFORMATION: potential protein kinase C, Prosite match  
US-09-338-907-4  
Query Match 100.0%; Score 1871; DB 4; Length 353;  
Best Local Similarity 100.0%; Pred. No. 1.9e-194;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRLPLSVVLLGTAPYYVLAWGVWRLSAFLPARFYQALDDRLCYVQSMVLEFFENYTG 60  
|||||  
DB 1 MRLPLSVVLLGTAPYYVLAWGVWRLSAFLPARFYQALDDRLCYVQSMVLEFFENYTG 60  
QY 61 VOILLYGDLPKKNENIYLANHOSYVDVADILAIQNALGHVRYVLEGGKMLPLYGC 120  
|||||  
DB 61 VOILLYGDLPKKNENIYLANHOSYVDVADILAIQNALGHVRYVLEGGKMLPLYGC 120  
QY 121 YFAOHGGIYVKSASAKNEKEMRNKLOSVDAGPMVLVFPBSTRNPEOTKYLASQAF 180  
|||||  
DB 121 YFAOHGGIYVKSASAKNEKEMRNKLOSVDAGPMVLVFPBSTRNPEOTKYLASQAF 180  
QY 181 AAQRLAVLKHVLTPIRIKATHVAFDCMKNYLDAIYDVTAVVYEGKDDGGORRESPTTEFL 240  
|||||  
DB 181 AAQRLAVLKHVLTPIRIKATHVAFDCMKNYLDAIYDVTAVVYEGKDDGGORRESPTTEFL 240  
QY 241 CKECPKIHIIHIDRIDKDDVPEDEHMRRLHEFFELKDKMLIEFYSPPERKRRPGKS 300  
|||||  
DB 241 CKECPKIHIIHIDRIDKDDVPEDEHMRRLHEFFELKDKMLIEFYSPPERKRRPGKS 300  
QY 301 VNSKLSIKRTPSMLTSLGTLGMLTDAGRKLYVTWITYGTLGLCMTYTIKA 353  
|||||  
DB 301 VNSKLSIKRTPSMLTSLGTLGMLTDAGRKLYVTWITYGTLGLCMTYTIKA 353  
RESULT 3

```

US-09-218-207-4
Sequence 4, Application US/09218207
Patent No. 6346381
GENERAL INFORMATION:
APPLICANT: Cohen, Daniel
APPLICANT: Blumenfeld, Marta
APPLICANT: Ilye, Chumakov
APPLICANT: Bougueleret, Lydie
TITLE OF INVENTION: Prostate cancer gene
FILE REFERENCE: GENSET_018CP1
CURRENT APPLICATION NUMBER: US/09/218.207
CURRENT FILING DATE: 1998-12-22
EARLIER APPLICATION NUMBER: 08/996,306
EARLIER FILING DATE: 1997-12-22
EARLIER APPLICATION NUMBER: 60/099,658
EARLIER FILING DATE: 1998-09-09
NUMBER OF SEQ ID NOS: 578
SOFTWARE: Patent.pm
SEQ ID NO 4
LENGTH: 353
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: HELIX
LOCATION: 1..33
OTHER INFORMATION: Rao and Argos identification method, potential helix
FEATURE:
NAME/KEY: HELIX
LOCATION: 4..20
OTHER INFORMATION: Klein, Kanehisa and Delisi identification method, potential helix
NAME/KEY: HELIX
LOCATION: 4..24
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, potential helix
FEATURE:
NAME/KEY: MYRISTATE
LOCATION: 12..16
OTHER INFORMATION: Prositte match
FEATURE:
NAME/KEY: HELIX
LOCATION: 50..70
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, potential helix
FEATURE:
NAME/KEY: CARBOHYD
LOCATION: 57..59
OTHER INFORMATION: Prositte match
FEATURE:
NAME/KEY: HELIX
LOCATION: 76..96
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, potential helix
FEATURE:
NAME/KEY: PHOSPHORYLATION
LOCATION: 84
OTHER INFORMATION: potential caseine kinase II site, Prositte match
FEATURE:
NAME/KEY: SITE
LOCATION: 94..115
OTHER INFORMATION: potential leucine zipper site, Prositte match
FEATURE:
NAME/KEY: MYRISTATE
LOCATION: 119..123
OTHER INFORMATION: potential site, Prositte match
FEATURE:
NAME/KEY: PHOSPHORYLATION
LOCATION: 133
OTHER INFORMATION: potential protein kinase C, Prositte match
FEATURE:
NAME/KEY: PHOSPHORYLATION
LOCATION: 147

```

	OTHER INFORMATION: potential casetine kinase II site, Prosite match	
/	FEATURE:	
/	NAME/KEY: PHOSPHORYLATION	
/	LOCATION: 194	
/	OTHER INFORMATION: potential protein kinase C, Prosite match	
/	FEATURE:	
/	NAME/KEY: PHOSPHORYLATION	
/	LOCATION: 215	
/	OTHER INFORMATION: potential Tyrosine kinase site, Prosite match	
/	FEATURE:	
/	NAME/KEY: SULFATATION	
/	LOCATION: 221	
/	OTHER INFORMATION: Prosite match	
/	FEATURE:	
/	NAME/KEY: PHOSPHORYLATION	
/	LOCATION: 233	
/	OTHER INFORMATION: potential CAMP and cGMP dependant protein kinase site,	
/	FEATURE:	
/	NAME/KEY: PHOSPHORYLATION	
/	LOCATION: 235	
/	OTHER INFORMATION: potential caseine kinase II site, Prosite match	
/	FEATURE:	
/	NAME/KEY: PHOSPHORYLATION	
/	LOCATION: 306	
/	OTHER INFORMATION: potential protein kinase C, Prosite match	
/	FEATURE:	
/	NAME/KEY: HELIX	
/	LOCATION: 310..330	
/	OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,	
/	FEATURE:	
/	NAME/KEY: MYRISTATE	
/	LOCATION: 319..323	
/	OTHER INFORMATION: Prosite match	
/	FEATURE:	
/	NAME/KEY: MYRISTATE	
/	LOCATION: 323..327	
/	OTHER INFORMATION: Prosite match	
/	FEATURE:	
/	NAME/KEY: AMIDATION	
/	LOCATION: 329	
/	OTHER INFORMATION: Prosite match	
/	FEATURE:	
/	NAME/KEY: HELIX	
/	LOCATION: 333..353	
/	OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,	
/	FEATURE:	
/	NAME/KEY: MYRISTATE	
/	LOCATION: 341..345	
/	OTHER INFORMATION: Prosite match	
/	FEATURE:	
/	NAME/KEY: PHOSPHORYLATION	
/	LOCATION: 350	
/	OTHER INFORMATION: potential protein kinase C, Prosite match	
/	US-09-218-207-4	
Query Match	100.0%; Score 1871; DB 4; Length 353;	
Best Local Similarity	100.0%; Pred. No.1,'9e-194;	
Matches	353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MRYLLPSVVLGTADPTYYLVANGVMRLSNFLPARFYQLDRLICYVSOMVLFPPENYTG 60	
DB	1 MRYLLPSVVLGTADPTYYLVANGVMRLSNFLPARFYQLDRLICYVSOMVLFPPENYTG 60	
QY	61 VOILLYGDLPRKKEINIYLANHQSIVDMIVADIILIRONALGHVRVYLKEGIKMLPLTGC 120	
DB	61 VOILLYGDLPRKKEINIYLANHQSIVDMIVADIILIRONALGHVRVYLKEGIKMLPLTGC 120	
QY	121 YFAOHGGIYVKRSAKFNKKEMRNKLQSYVDAGTPMYLVIIFPEGGTYNPDQTKVLSASOAF 180	
DB	121 YFAOHGGIYVKRSAKFNKKEMRNKLQSYVDAGTPMYLVIIFPEGGTYNPDQTKVLSASOAF 180	
QY	181 AAQRGLAVLKHLVPRIKATHVAFDCKMNYLDIAIVTVYVEGDKDGQGRRESPTMEFL 240	



Db 181 AAQRLAVLKHVLTPIRIKATHVAFDCMKNYLDAIDVTYVYEGKDDGGRRSPPTTEFL 240  
Qy 241 CKECKRIHIHDIRIDKDVPEEOEHMRRLHREFEIKDKMLIEFYESPDPERRKRPFGKS 300  
Db 241 CKECKRIHIHDIRIDKDVPEEOEHMRRLHREFEIKDKMLIEFYESPDPERRKRPFGKS 300  
Qy 301 VNSKLSIKKTLPSMLISGLTFAGMLMTDAGRKLVTYNTWITYGTLGCLMWTIKA 353  
Db 301 VNSKLSIKKTLPSMLISGLTFAGMLMTDAGRKLVTYNTWITYGTLGCLMWTIKA 353

RESULT 4  
US-08-996-306-5  
; Sequence 5, Application US/08996306  
; Patent No. 5945522

GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Chumakov, Ilya  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: Prostate cancer gene  
; NUMBER OF SEQUENCES: 68  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson & Bear  
; STREET: 501 West Broadway  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92101-3505  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Win95  
; SOFTWARE: Word  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/996.306  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israelson, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: GENSET.018A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 364 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: SINGLE  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: PROTEIN  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: potential protein  
; LOCATION: 1..364  
; US-08-996-306-5

Query Match 100.0%; Score 1871; DB 2; Length 364;  
Best Local Similarity 100.0%; Pred. No. 2e-194;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MRYLLPSVVLGTAPTYVLAMGVWRLLSAFLPARFYQALDDRLCYCYOSMVLFFPENYTG 60  
Db 12 MRYLLPSVVLGTAPTYVLAMGVWRLLSAFLPARFYQALDDRLCYCYOSMVLFFPENYTG 71  
Qy 61 VOILLYGDLPKKNENIITLANHOSYVDWIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120  
Db 72 VOILLYGDLPKKNENIITLANHOSYVDWIVADILAIRONALGHVRYVLEKGLKWLPLYGC 131

Qy 121 YFAOHGIIYVKSRAKFNEMKRNKLOSVDAGTAPMYLVTFPEGSTRYNPQOTVLSASQAF 180  
Db 132 YFAOHGIIYVKSRAKFNEMKRNKLOSVDAGTAPMYLVTFPEGSTRYNPQOTVLSASQAF 191  
Qy 181 AAQRLAVLKHVLTPIRIKATHVAFDCMKNYLDAIDVTYVYEGKDDGGRRSPPTTEFL 240  
Db 192 AAQRLAVLKHVLTPIRIKATHVAFDCMKNYLDAIDVTYVYEGKDDGGRRSPPTTEFL 251  
Qy 241 CKECKRIHIHDIRIDKDVPEEOEHMRRLHREFEIKDKMLIEFYESPDPERRKRPFGKS 300  
Db 252 CKECKRIHIHDIRIDKDVPEEOEHMRRLHREFEIKDKMLIEFYESPDPERRKRPFGKS 311  
Qy 301 VNSKLSIKKTLPSMLISGLTFAGMLMTDAGRKLVTYNTWITYGTLGCLMWTIKA 353  
Db 312 VNSKLSIKKTLPSMLISGLTFAGMLMTDAGRKLVTYNTWITYGTLGCLMWTIKA 364

RESULT 5  
US-09-338-907-5  
; Sequence 5, Application US/09338907  
; Patent No. 626546

GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Ilya, Chumakov  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: PROSTATE CANCER GENE  
; FILE REFERENCE: GENSET.18CP1CP  
; CURRENT APPLICATION NUMBER: US/09/338.907  
; EARLIER FILING DATE: 1999-06-23  
; EARLIER APPLICATION NUMBER: 08/996.306  
; EARLIER FILING DATE: 1997-12-22  
; EARLIER APPLICATION NUMBER: 60/099.658  
; EARLIER FILING DATE: 1998-09-09  
; EARLIER APPLICATION NUMBER: 09/218.207  
; EARLIER FILING DATE: 1998-12-22  
; NUMBER OF SEQ ID NOS: 578  
; SOFTWARE: Patent.pm  
; SEQ ID NO 5  
; LENGTH: 364  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-338-907-5

Query Match 100.0%; Score 1871; DB 4; Length 364;  
Best Local Similarity 100.0%; Pred. No. 2e-194;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRYLLPSVVLGTAPTYVLAMGVWRLLSAFLPARFYQALDDRLCYCYOSMVLFFPENYTG 60  
Db 12 MRYLLPSVVLGTAPTYVLAMGVWRLLSAFLPARFYQALDDRLCYCYOSMVLFFPENYTG 71  
Qy 61 VOILLYGDLPKKNENIITLANHOSYVDWIVADILAIRONALGHVRYVLEKGLKWLPLYGC 120  
Db 72 VOILLYGDLPKKNENIITLANHOSYVDWIVADILAIRONALGHVRYVLEKGLKWLPLYGC 131  
Qy 121 YFAOHGIIYVKSRAKFNEMKRNKLOSVDAGTAPMYLVTFPEGSTRYNPQOTVLSASQAF 180  
Db 132 YFAOHGIIYVKSRAKFNEMKRNKLOSVDAGTAPMYLVTFPEGSTRYNPQOTVLSASQAF 191  
Qy 181 AAQRLAVLKHVLTPIRIKATHVAFDCMKNYLDAIDVTYVYEGKDDGGRRSPPTTEFL 240  
Db 192 AAQRLAVLKHVLTPIRIKATHVAFDCMKNYLDAIDVTYVYEGKDDGGRRSPPTTEFL 251  
Qy 241 CKECKRIHIHDIRIDKDVPEEOEHMRRLHREFEIKDKMLIEFYESPDPERRKRPFGKS 300  
Db 252 CKECKRIHIHDIRIDKDVPEEOEHMRRLHREFEIKDKMLIEFYESPDPERRKRPFGKS 311  
Qy 301 VNSKLSIKKTLPSMLISGLTFAGMLMTDAGRKLVTYNTWITYGTLGCLMWTIKA 353  
Db 312 VNSKLSIKKTLPSMLISGLTFAGMLMTDAGRKLVTYNTWITYGTLGCLMWTIKA 364

```
RESULT 6
US-09-218-207-5
; Sequence 5, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; FILE REFERENCE: GENSET.018CPI
; CURRENT APPLICATION NUMBER: US/09/218, 207
; CURRENT FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 5
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-218-207-5
```

```
Query Match 100.0%; Score 1871; DB 4; Length 364;
Best Local Similarity 100.0%; Pred. No. 2e-194;
Matches 333; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MRYLLPSVLLGTAPTYVVLAMGWRLLSAFLPARFYQALDDRLCYVQSWLFEFFENYTG 60
DB 12 MRYLLPSVLLGTAPTYVVLAMGWRLLSAFLPARFYQALDDRLCYVQSWLFEFFENYTG 71
QY 61 VOILLGDDLPKKNENIITYLANHSTYDWMIVADILAIROMLGHVRYLKEGLMWLPYGC 120
DB 72 VOILLGDDLPKKNENIITYLANHSTYDWMIVADILAIROMLGHVRYLKEGLMWLPYGC 131
QY 121 YFAOHGIIYVRSKAKNEKEMRNKLOSVDAGTPMYLVIFPESTRYNPEOTKVLASQAF 180
DB 132 YFAOHGIIYVRSKAKNEKEMRNKLOSVDAGTPMYLVIFPESTRYNPEOTKVLASQAF 191
QY 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDGGQRESPTMEFL 240
DB 192 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDGGQRESPTMEFL 251
QY 241 CKCEPKIHIDRIKDDVPEBOEHMRMLHERFEIKDKMLIEFYSPPDERRRKRPFGKS 300
DB 252 CKCEPKIHIDRIKDDVPEBOEHMRMLHERFEIKDKMLIEFYSPPDERRRKRPFGKS 311
QY 301 VNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLVTWYIYGTLLGCLMWYIKA 353
DB 312 VNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLVTWYIYGTLLGCLMWYIKA 364
```

```
RESULT 7
US-09-338-907-134
; Sequence 134, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CPICP
; CURRENT APPLICATION NUMBER: US/09/338, 907
; CURRENT FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
```

```
; EARLIER APPLICATION NUMBER: 09/218, 207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 134
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 81..83
; OTHER INFORMATION: Box I
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 160..165
; OTHER INFORMATION: Box II
US-09-338-907-134
```

```
Query Match 87.9%; Score 1644; DB 4; Length 315;
Best Local Similarity 89.2%; Pred. No. 6.5e-170;
Matches 315; Conservative 0; Mismatches 0; Indels 38; Gaps 1;
```

```
QY 1 MRYLLPSVLLGTAPTYVVLAMGWRLLSAFLPARFYQALDDRLCYVQSWLFEFFENYTG 60
DB 1 MRYLLPSVLLGTAPTYVVLAMGWRLLSAFLPARFYQALDDRLCYVQSWLFEFFENYTG 60
QY 61 VOILLGDDLPKKNENIITYLANHSTYDWMIVADILAIROMLGHVRYLKEGLMWLPYGC 120
DB 61 VOILLGDDLPKKNENIITYLANHSTYDWMIVADILAIROMLGHVRYLKEGLMWLPYGC 120
QY 121 YFAOHGIIYVRSKAKNEKEMRNKLOSVDAGTPMYLVIFPESTRYNPEOTKVLASQAF 180
DB 121 YFAOHGIIYVRSKAKNEKEMRNKLOSVDAGTPMYLVIFPESTRYNPEOTKVLASQAF 180
QY 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDGGQRESPTMEFL 240
DB 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDGGQRESPTMEFL 202
QY 241 CKCEPKIHIDRIKDDVPEBOEHMRMLHERFEIKDKMLIEFYSPPDERRRKRPFGKS 300
DB 203 CKCEPKIHIDRIKDDVPEBOEHMRMLHERFEIKDKMLIEFYSPPDERRRKRPFGKS 262
QY 301 VNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLVTWYIYGTLLGCLMWYIKA 353
DB 263 VNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLVTWYIYGTLLGCLMWYIKA 315
```

```
RESULT 8
US-09-218-207-134
; Sequence 134, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; FILE REFERENCE: GENSET.018CPI
; CURRENT APPLICATION NUMBER: US/09/218, 207
; CURRENT FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 134
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
```

LOCATION: 81..83  
 OTHER INFORMATION: Box 1  
 FEATURE:  
 NAME/KEY: SITE  
 LOCATION: 160..165  
 OTHER INFORMATION: Box 11  
 US-09-218-207-134

Query Match 87.9%; Score 1644; DB 4; Length 315;  
 Best Local Similarity 89.2%; Pred. No. 6.5e-170; Indels 38; Gaps 1;  
 Matches 315; Conservative 0; Mismatches 0;

QY 1 MRYLLPSVLLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYVYOSVLFEEFNNTG 60  
 DB 1 MRYLLPSVLLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYVYOSVLFEEFNNTG 60  
 QY 61 VOILLYGDLPRKKNENIYLANHQSIVDWIVADILAIQNALGHVRYLKEGLKMLPLYGC 120  
 DB 61 VOILLYGDLPRKKNENIYLANHQSIVDWIVADILAIQNALGHVRYLKEGLKMLPLYGC 120  
 QY 121 YFAOHGIVYKRSKAFNEKEMRNKLOSIVYDAGTPMYLVIFPEGTRYNPQOTVYLSAQAF 180  
 DB 121 YFAOHGIVYKRSKAFNEKEMRNKLOSIVYDAGTPMYLVIFPEGTRYNPQOTVYLSAQAF 180  
 QY 181 AAORGLAVLKHVLTPIKATHVAFDCMKNYLDIAYDVTVVYEGKDDGGGRRSPMTTEF 240  
 DB 181 AAOR-----GKDDGGGRRSPMTTEF 202  
 QY 241 CKCEPKIHIDRIDKKVPEBOEHMRWLHERFEIKDKMLIEFYESPDPERRRRKFPCK 300  
 DB 203 CKCEPKIHIDRIDKKVPEBOEHMRWLHERFEIKDKMLIEFYESPDPERRRRKFPCK 262  
 QY 301 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTNTVYGTLLGCLWVTIKA 353  
 DB 263 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTNTVYGTLLGCLWVTIKA 315

# RESULT 9

US-09-338-907-74  
 Sequence 74, Application US/0938907  
 Patent No. 6265546  
 GENERAL INFORMATION:  
 APPLICANT: Cohen, Daniel  
 APPLICANT: Blumenfeld, Marta  
 APPLICANT: Ilya, Chumakov  
 APPLICANT: Bougueleret, Lydie  
 TITLE OF INVENTION: PROSTATE CANCER GENE  
 FILE REFERENCE: GENSET.18CPLCP  
 CURRENT APPLICATION NUMBER: US/09/338,907  
 EARLIER FILING DATE: 1998-06-23  
 EARLIER APPLICATION NUMBER: 08/996,306  
 EARLIER FILING DATE: 1997-12-22  
 EARLIER APPLICATION NUMBER: 60/099,658  
 EARLIER FILING DATE: 1998-09-09  
 EARLIER APPLICATION NUMBER: 09/218,207  
 EARLIER FILING DATE: 1998-12-22  
 NUMBER OF SEQ ID NOS: 578  
 SOFTWARE: Patent.pm  
 SEQ ID NO 74  
 LENGTH: 354  
 TYPE: PRT  
 ORGANISM: Mus musculus  
 US-09-338-907-74

Query Match 83.8%; Score 1568.5; DB 4; Length 354;  
 Best Local Similarity 81.4%; Pred. No. 1.2e-161; Indels 1; Gaps 1;  
 Matches 288; Conservative 39; Mismatches 26;

QY 1 MRYLLPSVLLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYVYOSVLFEEFNNTG 60  
 DB 1 MRYLLPSVLLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYVYOSVLFEEFNNTG 60

QY 61 VOILLYGDLPRKKNENIYLANHQSIVDWIVADILAIQNALGHVRYLKEGLKMLPLYGC 120  
 DB 61 VOILLYGDLPRKKNENIYLANHQSIVDWIVADILAIQNALGHVRYLKEGLKMLPLYGC 120  
 QY 121 YFAOHGIVYKRSKAFNEKEMRNKLOSIVYDAGTPMYLVIFPEGTRYNPQOTVYLSAQAF 180  
 DB 121 YFAOHGIVYKRSKAFNEKEMRNKLOSIVYDAGTPMYLVIFPEGTRYNPQOTVYLSAQAF 180  
 QY 181 AAORGLAVLKHVLTPIKATHVAFDCMKNYLDIAYDVTVVYEGKDDG-GGRRSPMTTEF 239  
 DB 181 AAORGLAVLKHVLTPIKATHVAFDCMKNYLDIAYDVTVVYEGKDDG-GGRRSPMTTEF 240  
 QY 240 CKCEPKIHIDRIDKKVPEBOEHMRWLHERFEIKDKMLIEFYESPDPERRRRKFPCK 299  
 DB 241 CKCEPKIHIDRIDKKVPEBOEHMRWLHERFEIKDKMLIEFYESPDPERRRRKFPCK 300  
 QY 300 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTNTVYGTLLGCLWVTIKA 353  
 DB 301 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTNTVYGTLLGCLWVTIKA 354

# RESULT 10

US-09-218-207-74  
 Sequence 74, Application US/09218207  
 Patent No. 6346381  
 GENERAL INFORMATION:  
 APPLICANT: Cohen, Daniel  
 APPLICANT: Blumenfeld, Marta  
 APPLICANT: Ilya, Chumakov  
 APPLICANT: Bougueleret, Lydie  
 TITLE OF INVENTION: Prostate cancer gene  
 FILE REFERENCE: GENSET.018CPL  
 CURRENT APPLICATION NUMBER: US/09/218,207  
 EARLIER FILING DATE: 1998-12-22  
 EARLIER APPLICATION NUMBER: 08/996,306  
 EARLIER FILING DATE: 1997-12-22  
 EARLIER APPLICATION NUMBER: 60/099,658  
 EARLIER FILING DATE: 1998-09-09  
 NUMBER OF SEQ ID NOS: 578  
 SOFTWARE: Patent.pm  
 SEQ ID NO 74  
 LENGTH: 354  
 TYPE: PRT  
 ORGANISM: Mus musculus  
 US-09-218-207-74

Query Match 83.8%; Score 1568.5; DB 4; Length 354;  
 Best Local Similarity 81.4%; Pred. No. 1.2e-161; Indels 1; Gaps 1;  
 Matches 288; Conservative 39; Mismatches 26;

QY 1 MRYLLPSVLLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYVYOSVLFEEFNNTG 60  
 DB 1 MRYLLPSVLLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYVYOSVLFEEFNNTG 60  
 QY 61 VOILLYGDLPRKKNENIYLANHQSIVDWIVADILAIQNALGHVRYLKEGLKMLPLYGC 120  
 DB 61 VOILLYGDLPRKKNENIYLANHQSIVDWIVADILAIQNALGHVRYLKEGLKMLPLYGC 120  
 QY 121 YFAOHGIVYKRSKAFNEKEMRNKLOSIVYDAGTPMYLVIFPEGTRYNPQOTVYLSAQAF 180  
 DB 121 YFAOHGIVYKRSKAFNEKEMRNKLOSIVYDAGTPMYLVIFPEGTRYNPQOTVYLSAQAF 180  
 QY 181 AAORGLAVLKHVLTPIKATHVAFDCMKNYLDIAYDVTVVYEGKDDG-GGRRSPMTTEF 239  
 DB 181 AAORGLAVLKHVLTPIKATHVAFDCMKNYLDIAYDVTVVYEGKDDG-GGRRSPMTTEF 240  
 QY 240 CKCEPKIHIDRIDKKVPEBOEHMRWLHERFEIKDKMLIEFYESPDPERRRRKFPCK 299  
 DB 241 CKCEPKIHIDRIDKKVPEBOEHMRWLHERFEIKDKMLIEFYESPDPERRRRKFPCK 300  
 QY 300 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTNTVYGTLLGCLWVTIKA 353

Db 301 SVHSRLSVKKTLPSSVLISGLTAVMLMTESGRKLYMGWTGLGCLMFWIKA 354

RESULT 11  
US-09-338-907-135  
Sequence 135, Application US/09338907  
Patent No. 6265546

GENERAL INFORMATION:

APPLICANT: Cohen, Daniel

APPLICANT: Blumenfeld, Marta

APPLICANT: Ilya, Chumakov

APPLICANT: Bougueleret, Lydie

TITLE OF INVENTION: PROSTATE CANCER GENE

FILE REFERENCE: GENSET.18CP1CP

CURRENT APPLICATION NUMBER: US/09/338,907

EARLIER FILING DATE: 1999-06-23

EARLIER APPLICATION NUMBER: 08/996,306

EARLIER FILING DATE: 1997-12-22

EARLIER APPLICATION NUMBER: 60/099,658

EARLIER FILING DATE: 1998-09-09

EARLIER APPLICATION NUMBER: 09/218,207

EARLIER FILING DATE: 1998-12-22

NUMBER OF SEQ ID NOS: 578

SOFTWARE: Patent.pm

SEQ ID NO 135

LENGTH: 300

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SITE

LOCATION: 81..83

OTHER INFORMATION: Box I

NAME/KEY: SITE

LOCATION: 160..165

OTHER INFORMATION: Box II

US-09-338-907-135

Query Match 83.1%; Score 1555.5; DB 4; Length 300;  
Best Local Similarity 85.0%; Pred. No. 2.4e-160;  
Matches 300; Conservative 0; Mismatches 0; Indels 53; Gaps 1;

QY 1 MRYLPSVVLGTAPTYVLAAGVWRLLSAFLPARFYQALDDRLCYCYOSVWLFEEFNYTG 60  
Db 1 MRYLPSVVLGTAPTYVLAAGVWRLLSAFLPARFYQALDDRLCYCYOSVWLFEEFNYTG 60  
QY 61 VOILLYGDLPKKNENITYLANHOSYVDWIVADILAIROMALGHVRYVLKGLKWLPLYGC 120  
Db 61 VOILLYGDLPKKNENITYLANHOSYVDWIVADILAIROMALGHVRYVLKGLKWLPLYGC 120  
QY 121 YFAOHGITYVRSKAKNEKEMRNKLOSVDAGTPLYVTFPESTRYNPEOTKVLASQAF 180  
Db 121 YFAOHGITYVRSKAKNEKEMRNKLOSVDAGTPLYVTFPESTRYNPEOTKVLASQAF 180  
QY 181 AAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVYEGKDGQGRRESPTMTEFL 240  
Db 181 AAOR-----EFL 187  
QY 241 CKECPRIHIHIDRIDKDVPEOEHRMRLHERFEIKDKMLIEFYSPDERRRKRPFGKS 300  
Db 188 CKECPRIHIHIDRIDKDVPEOEHRMRLHERFEIKDKMLIEFYSPDERRRKRPFGKS 247  
QY 301 VNSKLSIKKTLPSSVLISGLTAGMLMTDAGRKLYVNTWYIGTLLGLMWTIKA 353  
Db 248 VNSKLSIKKTLPSSVLISGLTAGMLMTDAGRKLYVNTWYIGTLLGLMWTIKA 300

RESULT 12  
US-09-218-207-135  
Sequence 135, Application US/09218207  
Patent No. 6346381

GENERAL INFORMATION:

APPLICANT: Cohen, Daniel

APPLICANT: Blumenfeld, Marta

APPLICANT: Ilya, Chumakov

TITLE OF INVENTION: Prostate cancer gene

FILE REFERENCE: GENSET.018CP1

CURRENT APPLICATION NUMBER: US/09/218,207

EARLIER FILING DATE: 1998-12-22

EARLIER APPLICATION NUMBER: 08/996,306

EARLIER FILING DATE: 1997-12-22

EARLIER APPLICATION NUMBER: 60/099,658

EARLIER FILING DATE: 1998-09-09

NUMBER OF SEQ ID NOS: 578

SOFTWARE: Patent.pm

SEQ ID NO 135

LENGTH: 300

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SITE

LOCATION: 81..83

OTHER INFORMATION: Box I

NAME/KEY: SITE

LOCATION: 160..165

OTHER INFORMATION: Box II

US-09-218-207-135

Query Match 83.1%; Score 1555.5; DB 4; Length 300;  
Best Local Similarity 85.0%; Pred. No. 2.4e-160;  
Matches 300; Conservative 0; Mismatches 0; Indels 53; Gaps 1;

QY 1 MRYLPSVVLGTAPTYVLAAGVWRLLSAFLPARFYQALDDRLCYCYOSVWLFEEFNYTG 60  
Db 1 MRYLPSVVLGTAPTYVLAAGVWRLLSAFLPARFYQALDDRLCYCYOSVWLFEEFNYTG 60  
QY 61 VOILLYGDLPKKNENITYLANHOSYVDWIVADILAIROMALGHVRYVLKGLKWLPLYGC 120  
Db 61 VOILLYGDLPKKNENITYLANHOSYVDWIVADILAIROMALGHVRYVLKGLKWLPLYGC 120  
QY 121 YFAOHGITYVRSKAKNEKEMRNKLOSVDAGTPLYVTFPESTRYNPEOTKVLASQAF 180  
Db 121 YFAOHGITYVRSKAKNEKEMRNKLOSVDAGTPLYVTFPESTRYNPEOTKVLASQAF 180  
QY 181 AAORGLAVLKHVLTLPRIKATHVAFDCMKNYLDAIYDVTVYEGKDGQGRRESPTMTEFL 240  
Db 181 AAOR-----EFL 187  
QY 241 CKECPRIHIHIDRIDKDVPEOEHRMRLHERFEIKDKMLIEFYSPDERRRKRPFGKS 300  
Db 188 CKECPRIHIHIDRIDKDVPEOEHRMRLHERFEIKDKMLIEFYSPDERRRKRPFGKS 247  
QY 301 VNSKLSIKKTLPSSVLISGLTAGMLMTDAGRKLYVNTWYIGTLLGLMWTIKA 353  
Db 248 VNSKLSIKKTLPSSVLISGLTAGMLMTDAGRKLYVNTWYIGTLLGLMWTIKA 300

RESULT 13  
US-09-338-907-127  
Sequence 127, Application US/09338907  
Patent No. 6265546  
GENERAL INFORMATION:  
APPLICANT: Cohen, Daniel  
APPLICANT: Blumenfeld, Marta  
APPLICANT: Ilya, Chumakov  
APPLICANT: Bougueleret, Lydie  
TITLE OF INVENTION: PROSTATE CANCER GENE  
FILE REFERENCE: GENSET.18CP1CP  
CURRENT APPLICATION NUMBER: US/09/338,907  
EARLIER FILING DATE: 1999-06-23  
EARLIER APPLICATION NUMBER: 08/996,306

```

; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218,207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 127
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 98..103
; OTHER INFORMATION: Box II
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 149..157
; OTHER INFORMATION: Box III
US-09-338-907-127
```

```

Query Match      80.1%; Score 1498; DB 4; Length 291;
Best Local Similarity 82.4%; Pred. No. 3.9e-154;
Matches 291; Conservative 0; Mismatches 0; Indels 62; Gaps 1;

QY 1 MRYLLPSVLLGTAPTYVLA MGWRLLSAFLPARFYQALDDRLCYQSVLFFFNNTG 60
    |||||||
DB 1 MRYLLPSVLLGTAPTYVLA MGWRLLSAFLPARFYQALDDRLCYQSVLFFFNNTG 60
QY 61 VOILLYGDLPKKNENIYYLANHOSYDWMIVADILAIQNALGHVRYLKGKMLPLYGC 120
    |
DB 61 V-----
QY 121 YFAOHGIIYKRSKAKNEKEMRNKLOSVDAGTPMYLVIFPEGT RYNPEOTKVLASQAF 180
    |||||||
DB 62 ---QHGGIYKRSKAKNEKEMRNKLOSVDAGTPMYLVIFPEGT RYNPEOTKVLASQAF 118
QY 181 AAOAGLAVLKHYLTPRIKATHVAFDCMKNYLDAIYDVTYVVEGKDGGQRESPTMTEFL 240
    |||||||
DB 119 AAOAGLAVLKHYLTPRIKATHVAFDCMKNYLDAIYDVTYVVEGKDGGQRESPTMTEFL 178
QY 241 CKCEPKIHIDRIDKDDVEEOEHMRMLHERFEIKDKMLIEFYSPDERKRKRPFGKS 300
    |||||||
DB 179 CKCEPKIHIDRIDKDDVEEOEHMRMLHERFEIKDKMLIEFYSPDERKRKRPFGKS 238
QY 301 VNSKLSIKKTLP SMLILSGLTAGMLMTDAGRKLYVNTWIYGTLLGCLMTYIKA 353
    |||||||
DB 239 VNSKLSIKKTLP SMLILSGLTAGMLMTDAGRKLYVNTWIYGTLLGCLMTYIKA 291
```

```

RESULT 14
US-09-218-207-127
; Sequence 127, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; FILE REFERENCE: GENSET.018CPI
; CURRENT APPLICATION NUMBER: US/09/218,207
; CURRENT FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 127
; LENGTH: 291
; TYPE: PRT
```

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 98..103
; OTHER INFORMATION: Box II
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 149..157
; OTHER INFORMATION: Box III
US-09-218-207-127
```

```

Query Match      80.1%; Score 1498; DB 4; Length 291;
Best Local Similarity 82.4%; Pred. No. 3.9e-154;
Matches 291; Conservative 0; Mismatches 0; Indels 62; Gaps 1;

QY 1 MRYLLPSVLLGTAPTYVLA MGWRLLSAFLPARFYQALDDRLCYQSVLFFFNNTG 60
    |||||||
DB 1 MRYLLPSVLLGTAPTYVLA MGWRLLSAFLPARFYQALDDRLCYQSVLFFFNNTG 60
QY 61 VOILLYGDLPKKNENIYYLANHOSYDWMIVADILAIQNALGHVRYLKGKMLPLYGC 120
    |
DB 61 V-----
QY 121 YFAOHGIIYKRSKAKNEKEMRNKLOSVDAGTPMYLVIFPEGT RYNPEOTKVLASQAF 180
    |||||||
DB 62 ---QHGGIYKRSKAKNEKEMRNKLOSVDAGTPMYLVIFPEGT RYNPEOTKVLASQAF 118
QY 181 AAOAGLAVLKHYLTPRIKATHVAFDCMKNYLDAIYDVTYVVEGKDGGQRESPTMTEFL 240
    |||||||
DB 119 AAOAGLAVLKHYLTPRIKATHVAFDCMKNYLDAIYDVTYVVEGKDGGQRESPTMTEFL 178
QY 241 CKCEPKIHIDRIDKDDVEEOEHMRMLHERFEIKDKMLIEFYSPDERKRKRPFGKS 300
    |||||||
DB 179 CKCEPKIHIDRIDKDDVEEOEHMRMLHERFEIKDKMLIEFYSPDERKRKRPFGKS 238
QY 301 VNSKLSIKKTLP SMLILSGLTAGMLMTDAGRKLYVNTWIYGTLLGCLMTYIKA 353
    |||||||
DB 239 VNSKLSIKKTLP SMLILSGLTAGMLMTDAGRKLYVNTWIYGTLLGCLMTYIKA 291
```

```

RESULT 15
US-09-338-907-128
; Sequence 128, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CPICP
; CURRENT APPLICATION NUMBER: US/09/338,907
; CURRENT FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218,207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 128
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 68..73
; OTHER INFORMATION: Box II
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 119..127
```

; OTHER INFORMATION: Box III  
US-09-338-907-128

Query Match 70.8%; Score 1325; DB 4; Length 261;  
Best Local Similarity 73.9%; Pred. NO. 1.9e-135;  
Matches 261; Conservative 0; Mismatches 0; Indels 92; Gaps 1;

```
QY 1 MRYLLPSVVLGTAPTYVVLANGVWRLLSAFLPARFYQALDDRLCYVQSMVLFEEENTYG 60
    |||||||
Db 1 MRYLLPSVVLGTAPTYVVLANGVWRLLSAFLPARFYQALDDRLCYVQSMVLFEEENTYG 60
QY 61 VQILLYGDLPKNKENIYYLANHSTVDWIVADILAIRQNALGHVRYVLEKGLKMLPLYGC 120
    ||
Db 61 VQ----- 62
QY 121 YFAQHGGIYVKRSKAFNEKEMRNKLQSYVDAGTPMYLYIPEEGTRYNPEOTKVLASQAF 180
    |||||||
Db 63 -----MYLYIPEEGTRYNPEOTKVLASQAF 88
QY 181 AAQGLAVLKHVLPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDGQRRSEPTMTPEFL 240
    |||||||
Db 89 AAQGLAVLKHVLPRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDGQRRSEPTMTPEFL 148
QY 241 CKECPKIHIDRIDKDKDVPPEQEHMRRLHEREIKDKMLIEFYESPDPERRKRFPGKS 300
    |||||||
Db 149 CKECPKIHIDRIDKDKDVPPEQEHMRRLHEREIKDKMLIEFYESPDPERRKRFPGKS 208
QY 301 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTNTWYGTGLGCLMVTIKA 353
    |||||||
Db 209 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTNTWYGTGLGCLMVTIKA 261
```

Search completed: August 28, 2002, 11:08:10  
Job time: 289 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:06:01 ; Search time 306.96 Seconds  
(without alignments)  
404.773 Million cell updates/sec

Title: US-09-853-526-4

Sequence: 1 MRYLLPSVLLGTAFTYVLA.....YVNTWITGLICLWVTKA 353

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 3502263 seqs, 351980561 residues

Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main:\*

1: /cgn2\_6/ptodata/2/paa/PCUS\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep:\*  
7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep:\*  
8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep:\*  
9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep:\*  
10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep:\*  
11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep:\*  
12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep:\*  
13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep:\*  
14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep:\*  
15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep:\*  
16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep:\*  
17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep:\*  
18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep:\*  
19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep:\*  
20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep:\*  
21: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep:\*  
22: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep:\*  
23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep:\*  
24: /cgn2\_6/ptodata/2/paa/US100\_COMB.pep:\*  
25: /cgn2\_6/ptodata/2/paa/US101\_COMB.pep:\*  
26: /cgn2\_6/ptodata/2/paa/US60\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1871	100.0	353	1	PCT-US01-01431-59
2	1871	100.0	353	1	PCT-US01-01431-79
3	1871	100.0	353	1	PCT-US01-11988-1474
4	1871	100.0	353	1	PCT-US01-11988-1475
5	1871	100.0	353	22	US-09-833-245-1474
6	1871	100.0	353	22	US-09-833-245-1475
7	1871	100.0	353	22	US-09-853-526-4

8	1871	100.0	353	23	US-09-901-484A-4	Sequence 4, Appl
9	1871	100.0	353	23	US-09-915-582-59	Sequence 59, Appl
10	1871	100.0	353	23	US-09-915-582-79	Sequence 79, Appl
11	1871	100.0	353	26	US-60-099-658-4	Sequence 4, Appl
12	1871	100.0	364	18	US-09-436-919-1	Sequence 7, Appl
13	1871	100.0	364	22	US-09-817-910-7	Sequence 1, Appl
14	1871	100.0	364	22	US-09-853-526-5	Sequence 5, Appl
15	1871	100.0	364	23	US-09-901-484A-5	Sequence 5, Appl
16	1871	100.0	364	26	US-60-099-658-5	Sequence 5, Appl
17	1861	99.5	353	18	US-09-488-725A-2736	Sequence 2736, Ap
18	1860	99.4	372	18	US-09-488-725A-6308	Sequence 6308, Ap
19	1644	87.9	315	22	US-09-853-526-134	Sequence 134, App
20	1644	87.9	315	23	US-09-901-484A-134	Sequence 110, App
21	1619.5	86.6	450	26	US-60-245-222-110	Sequence 74, Appl
22	1568.5	83.8	354	22	US-09-853-526-74	Sequence 74, Appl
23	1568.5	83.8	354	23	US-09-901-484A-74	Sequence 74, Appl
24	1568.5	83.8	354	26	US-60-099-658-74	Sequence 309, App
25	1559.5	83.4	1032	26	US-60-212-413-309	Sequence 248, App
26	1559.5	83.4	1032	26	US-60-229-518-248	Sequence 135, App
27	1555.5	83.1	300	22	US-09-853-526-135	Sequence 135, App
28	1555.5	83.1	300	23	US-09-901-484A-135	Sequence 127, App
29	1498	80.1	291	22	US-09-853-526-127	Sequence 127, App
30	1498	80.1	291	23	US-09-901-484A-127	Sequence 47, Appl
31	1427	76.3	269	1	PCT-US01-01327-47	Sequence 128, App
32	1325	70.8	261	22	US-09-853-526-128	Sequence 128, App
33	1325	70.8	261	23	US-09-901-484A-128	Sequence 70, Appl
34	1203	64.3	228	22	US-09-853-526-70	Sequence 70, Appl
35	1203	64.3	228	23	US-09-901-484A-70	Sequence 126, App
36	1203	64.3	228	26	US-60-099-658-70	Sequence 126, App
37	1182.5	63.2	238	22	US-09-853-526-126	Sequence 136, App
38	1182.5	63.2	238	23	US-09-901-484A-126	Sequence 136, App
39	976	52.2	185	22	US-09-853-526-136	Sequence 136, App
40	976	52.2	185	23	US-09-901-484A-136	Sequence 5, Appl
41	936	50.1	176	18	US-09-436-919-5	Sequence 133, App
42	820	43.8	182	22	US-09-853-526-133	Sequence 133, App
43	820	43.8	182	23	US-09-901-484A-133	Sequence 132, App
44	449	24.0	97	22	US-09-853-526-132	
45	449	24.0	97	23	US-09-901-484A-132	

#### ALIGNMENTS

RESULT 1  
PCT-US01-01431-59  
Sequence 59, Application PC/TUS0101431  
GENERAL INFORMATION:  
APPLICANT: Human Genome Sciences, Inc.  
TITLE OF INVENTION: 17 human secreted proteins  
FILE REFERENCE: P5723PCT  
CURRENT APPLICATION NUMBER: PCT/US01/01431  
CURRENT FILING DATE: 2001-01-17  
PRIOR APPLICATION NUMBER: 60/179,065  
PRIOR FILING DATE: 2000-01-31  
PRIOR APPLICATION NUMBER: 60/180,628  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: 60/231,968  
PRIOR FILING DATE: 2000-09-12  
NUMBER OF SEQ ID NOS: 97  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 59  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
PCT-US01-01431-59

Query Match 100.0%; Score 1871; DB 1; Length 353;  
Best Local Similarity 100.0%; Pred. No. 4.1e-184;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 MRYLLPSVLLGTAFTYVLAAGVWRLSAFLPRFQALDDRLCYVQSMVLEFFENYTG 60  
|||||

```
Db 1 MRYLLPSVLLGTAPTYVLAWGVWRLLSAFLPARFYQALDRLRYCYQSVWLEFFENYTG 60
QY 61 VOILLYGDLPKKNENIYLANHSTVDMIVADILAIROMALGHVRYVLEKGLKMLPLXGC 120
Db 61 VOILLYGDLPKKNENIYLANHSTVDMIVADILAIROMALGHVRYVLEKGLKMLPLXGC 120
QY 121 YFAOHGIIYVRSKAFNEKEKRNKLOSVDAGTPMYLVIPPEGTRYNPEQTKVLSASQAF 180
Db 121 YFAOHGIIYVRSKAFNEKEKRNKLOSVDAGTPMYLVIPPEGTRYNPEQTKVLSASQAF 180
QY 181 AAORGGLAVLKHLVLPRIKATHAVAFDCKMKNYLDALYDVTVVYEGKDDGQRRSEPTMTTEFL 240
Db 181 AAORGGLAVLKHLVLPRIKATHAVAFDCKMKNYLDALYDVTVVYEGKDDGQRRSEPTMTTEFL 240
QY 241 CKECPKIHIDRIDDKDVPEDQEHMRRWLHEREIKDKMLIEFYESPDPERRRRPFPGKS 300
Db 241 CKECPKIHIDRIDDKDVPEDQEHMRRWLHEREIKDKMLIEFYESPDPERRRRPFPGKS 300
QY 301 VNSKLSIKTKLPSMLILSGLTAGMLMTDAGRKLVTNTWYIGTLGCLMWITKA 353
Db 301 VNSKLSIKTKLPSMLILSGLTAGMLMTDAGRKLVTNTWYIGTLGCLMWITKA 353
```

```
RESULT 2
PCT-US01-01431-79
; Sequence 79, Application PC/TUS0101431
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: 17 human secreted proteins
; FILE REFERENCE: PS723PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01431
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/231,968
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 79
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-01431-79
```

```
Query Match 100.0%; Score 1871; DB 1; Length 353;
Best Local Similarity 100.0%; Pred. No. 4,1e-184;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MRYLLPSVLLGTAPTYVLAWGVWRLLSAFLPARFYQALDRLRYCYQSVWLEFFENYTG 60
Db 1 MRYLLPSVLLGTAPTYVLAWGVWRLLSAFLPARFYQALDRLRYCYQSVWLEFFENYTG 60
QY 61 VOILLYGDLPKKNENIYLANHSTVDMIVADILAIROMALGHVRYVLEKGLKMLPLXGC 120
Db 61 VOILLYGDLPKKNENIYLANHSTVDMIVADILAIROMALGHVRYVLEKGLKMLPLXGC 120
QY 121 YFAOHGIIYVRSKAFNEKEKRNKLOSVDAGTPMYLVIPPEGTRYNPEQTKVLSASQAF 180
Db 121 YFAOHGIIYVRSKAFNEKEKRNKLOSVDAGTPMYLVIPPEGTRYNPEQTKVLSASQAF 180
QY 181 AAORGGLAVLKHLVLPRIKATHAVAFDCKMKNYLDALYDVTVVYEGKDDGQRRSEPTMTTEFL 240
Db 181 AAORGGLAVLKHLVLPRIKATHAVAFDCKMKNYLDALYDVTVVYEGKDDGQRRSEPTMTTEFL 240
QY 241 CKECPKIHIDRIDDKDVPEDQEHMRRWLHEREIKDKMLIEFYESPDPERRRRPFPGKS 300
Db 241 CKECPKIHIDRIDDKDVPEDQEHMRRWLHEREIKDKMLIEFYESPDPERRRRPFPGKS 300
QY 301 VNSKLSIKTKLPSMLILSGLTAGMLMTDAGRKLVTNTWYIGTLGCLMWITKA 353
Db 301 VNSKLSIKTKLPSMLILSGLTAGMLMTDAGRKLVTNTWYIGTLGCLMWITKA 353
```

```
Db 301 VNSKLSIKTKLPSMLILSGLTAGMLMTDAGRKLVTNTWYIGTLGCLMWITKA 353
RESULT 3
PCT-US01-11988-1474
; Sequence 1474, Application PC/TUS0111988
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: PCT/US01/11988
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1474
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-11988-1474
```

```
Query Match 100.0%; Score 1871; DB 1; Length 353;
Best Local Similarity 100.0%; Pred. No. 4,1e-184;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MRYLLPSVLLGTAPTYVLAWGVWRLLSAFLPARFYQALDRLRYCYQSVWLEFFENYTG 60
Db 1 MRYLLPSVLLGTAPTYVLAWGVWRLLSAFLPARFYQALDRLRYCYQSVWLEFFENYTG 60
QY 61 VOILLYGDLPKKNENIYLANHSTVDMIVADILAIROMALGHVRYVLEKGLKMLPLXGC 120
Db 61 VOILLYGDLPKKNENIYLANHSTVDMIVADILAIROMALGHVRYVLEKGLKMLPLXGC 120
QY 121 YFAOHGIIYVRSKAFNEKEKRNKLOSVDAGTPMYLVIPPEGTRYNPEQTKVLSASQAF 180
Db 121 YFAOHGIIYVRSKAFNEKEKRNKLOSVDAGTPMYLVIPPEGTRYNPEQTKVLSASQAF 180
QY 181 AAORGGLAVLKHLVLPRIKATHAVAFDCKMKNYLDALYDVTVVYEGKDDGQRRSEPTMTTEFL 240
Db 181 AAORGGLAVLKHLVLPRIKATHAVAFDCKMKNYLDALYDVTVVYEGKDDGQRRSEPTMTTEFL 240
QY 241 CKECPKIHIDRIDDKDVPEDQEHMRRWLHEREIKDKMLIEFYESPDPERRRRPFPGKS 300
Db 241 CKECPKIHIDRIDDKDVPEDQEHMRRWLHEREIKDKMLIEFYESPDPERRRRPFPGKS 300
QY 301 VNSKLSIKTKLPSMLILSGLTAGMLMTDAGRKLVTNTWYIGTLGCLMWITKA 353
Db 301 VNSKLSIKTKLPSMLILSGLTAGMLMTDAGRKLVTNTWYIGTLGCLMWITKA 353
```

```
RESULT 4
PCT-US01-11988-1475
; Sequence 1475, Application PC/TUS0111988
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: PCT/US01/11988
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
```



```

; SEQ ID NO 1475
; LENGTH: 353
; TYPE: prt
; ORGANISM: Homo sapiens
PCT-US01-11988-1475

```

Query Match	100.0%;	Score 1871;	DB 1;	Length 353;
Best Local Similarity	100.0%;	Pred. No. 4.1e-184;		
Matches 353;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

**OY** 1 MRYLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDRLCYVQSVMLEFFENYTG 60  
**Db** 1 MRYLLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDRLCYVQSVMLEFFENYTG 60

OY 61 VQILLYGDLPRKNKENTIIYLANHQSTVDIMIVADILAIRONALGHVRYVLKSGELKWLEPLYGC 120  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
DB 61 VQILLYGDLPRKNKENTIIYLANHQSTVDIMIVADILAIRONALGHVRYVLKSGELKWLEPLYGC 120

[illegible]

**OY**    181 AAGGGLAVLKHLVTPRIKATHAFDCKMKNYLDAIYDVTVVEGGDDGGQRRSPMTTEFL 240  
     |||||  
**Db**    181 AAORGGLAVLKHVLTPRIKATHAFDCKMKNYLDAIYDVTVVEEGDDGGQRRSPMTTEFL 240  
     |||||

```

QY 241 CKEPKIHIDRIDKKDVPPEQEHMRWLHERFEIKDKMLIEFESPDDERRKKRPKGS 3000
      |||||
Db 241 CKECPKIHIDRIDKKDVPPEQEHMRWLHERFEIKDKMLIEFESPDDERRKKRPKGS 3000

```

QY 301 VNSKLSIKKTLPSMLILSGITAGMLMDAGRKLKYVNTWITYGTLGCLMWTIKA 353  
|||||  
Db 301 VNSKLSIKKTLPSMLILSGITAGMLMDAGRKLKYVNTWITYGTLGCLMWTIKA 353

```

RESULT      5
US-09-833-245-1474
; Sequence 1474, Application US/09833245
; GENERAL INFORMATION:

```

```

1  APPLICANT: Human Genome Sciences, Inc.
2  TITLE OF INVENTION: Albumin Fusion Proteins
3  FILE REFERENCE: PF546PCr
4  CURRENT APPLICATION NUMBER: US/09/833,245
5  CURRENT FILING DATE: 2001-04-12
6  PRIOR APPLICATION NUMBER: 60/229, 358
7  PRIOR FILING DATE: 2000-04-12
8  PRIOR APPLICATION NUMBER: 60/256, 931
9  PRIOR FILING DATE: 2000-12-21
10 PRIOR APPLICATION NUMBER: 60/199, 384
11 PRIOR FILING DATE: 2000-04-25
12 NUMBER OF SEQ ID NOS: 2267
13 SOFTWARE: Patentl Ver. 2.1
14 SEQ ID NO: 1474
15 LENGTH: 353
16 TYPE: prt
17 ORGANISM: Homo sapiens
18 OS-09-833-245-1474

```

Query Match	100.0%;	Score 1871;	DB 22;	Length 353;
Best Local Similarity	100.0%;	Pred. No. 4.1e-184;		
Matches 353;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 1 MRYLLPSVLLGTAPTYVLAMGWRLLSAFLPARFYQALDRLCYVQSMVLEFFENYTG 60  
|||||  
1 MRYLLPSVLLGTAPTYVLAMGWRLLSAFLPARFYQALDRLCYVQSMVLEFFENYTG 60

Qy 61 VQIIILYGDLEPKNKENIYYLANHQSTVDIMVADILAIRONALGHRYVLKESLKWPLELYGC 120  
Db 61 VQIIILYGDLEPKNKENIYYLANHQSTVDIMVADILAIRONALGHRYVLKESLKWPLELYGC 120

121 YFAQHGGLYKRSAFNEKEEMRNKLQSYVDAGIPMLVLFPEGIRYNPEQIRKVLASQAF 180

Db 121 YFAQHGGIYARSAKNEKEMERKNLQSTYDAGTPIVYIPEEGTRINPEOTKLASQAF 180

QY 181 AAORGIAVLKIVHLVPRIKATIVAFDCNNKYNLDAYDVTVVYEGKDGDGGQRRSPYTMTEFL 240

Db 181 AAQKGLAVLKIVHLVPRIKATIVAFDCNNKYNLDAYDVTVVYEGKDGDGGQRRSPYTMTEFL 240

QY 241 CKECPKIHIIHIDRIDDKDVPDEQEHMRMWLHEREIINDKMLIEFYESPDERRRKRRPPGKS 300

Db 241 CKECKKIHIIHIDRIDDKDVPDEQEHMRMWLHEREIINDKMLIEFYESPDERRRKRRPPGKS 300

QY 301 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLYVNTWIVYIGTLGLGLMTVITKA 353

Db 301 VNSKLSIKKTLPSMLILSGLTAGMLMTDAGRKLYVNTWIVYIGTLGLGLMTVITKA 353

RESULT 6  
US-09-833-245-1475  
Sequence 1475 Application IIS/09833245

```

1  GENERAL INFORMATION:
2  APPLICANT: Human Genome Sciences, Inc.
3  TITLE OF INVENTION: Albumin Fusion Proteins
4  FILE REFERENCE: PF546PCT
5  CURRENT APPLICATION NUMBER: US/09/833,245
6  CURRENT FILING DATE: 2001-04-12
7  PRIOR APPLICATION NUMBER: 60/229, 358
8  PRIOR FILING DATE: 2000-04-12
9  PRIOR APPLICATION NUMBER: 60/256, 931
10 PRIOR FILING DATE: 2000-12-21
11 PRIOR APPLICATION NUMBER: 60/199, 384
12 PRIOR FILING DATE: 2000-04-25
13 NUMBER OF SEQ ID NOS: 2267
14 SOFTWARE: PatentIn Ver. 2.1
15 SEQ ID NO 1475
16 LENGTH: 353
17 TYPE: prt
18 ORGANISM: Homo sapiens
19 US-09-833-245-1475

```

Query Match	100.0%;	Score 1871;	DB 22;	Length 353;
Best Local Similarity	100.0%;	Pred. No. 4.1e-184;		
Matches 353;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

**QY** 1 MRYLLPSVVLGTAPFYVLAWGVRLISAFLPARFYQLDDRLCYQSNAFLFEENNTG 60  
**Db** 1 MRYLPPSVVLGTAPFYVLAWGVRLISAFLPARFYQLDDRCLCVYSNVLFFENNVTG 60			

QY 61 VQILLGDEPKNKENTYYLANHQSTVMDIVADILAIRQNALGHRVYLKELKWLPLEYGC 120  
 |||||  
 DB 61 VQILLGDEPKNKENTYYLANHQSTVMDIVADILAIRQNALGHRVYLKELKWLPLEYGC 120

**QY** 121 YFAHGGIYKRSAKNEKEMNKLQSVDAGTPLYIFPEGTRYNPEDQTVLASQA F 180  
**Db** 121 YFAHGGIYKRSAKNEKEMNKLQSVDAGTPLYIFPEGTRYNPEDQTVLASQA F 180

QY 181 AAQRGLAVLKHVLTPIKATHAFAFCMKNYLDIAYDVIVVEGADGGGRRESPTWTEFL 240  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
DB 181 AAQRGLAVLKHVLTPIKATHAFAFCMKNYLDIAYDVIVVEGADGGGRRESPTWTEFL 240

QY 241 CKECKIHIDRIDKDDVPEQEHMRMLHERFEIKDMLIEFESPDERRKKRPGKS 300  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
DB 241 CKECPKIHIDRIDKDDVPEQHMRRLHERFEIKDKMLIEFYSPDERRRKRPPGKS 300

QY 301 VNSKLSIKKTTESMLILSGLTAGMLMTDAGRKLYVNNTWITYGTLGGLMVTIKA 353  
|||||  
Db 301 VNSKLSIKKTTESMLILSGLTAGMLMTDAGRKLYVNNTWITYGTLGGLMVTIKA 353

RESULT 7  
US-09-853-526-4  
; Sequence 4, Application US/09853526  
; GENERAL INFORMATION:

APPLICANT: Cohen, Daniel  
APPLICANT: Blumenfeld, Marta  
APPLICANT: Ilya, Chumakov  
TITLE OF INVENTION: PROSTATE CANCER GENE  
FILE REFERENCE: GENSET.18CPICP  
CURRENT APPLICATION NUMBER: US/09/853,526  
PRIORITY FILING DATE: 2001-05-11  
PRIORITY APPLICATION NUMBER: 09/338,907  
PRIORITY FILING DATE: 1999-06-23  
PRIORITY APPLICATION NUMBER: 08/996,306  
PRIORITY FILING DATE: 1997-12-22  
PRIORITY APPLICATION NUMBER: 60/099,658  
PRIORITY FILING DATE: 1998-09-09  
PRIORITY APPLICATION NUMBER: 09/218,207  
PRIORITY FILING DATE: 1998-12-22  
NUMBER OF SEQ ID NOS: 578  
SOFTWARE: Patent.pm  
SEQ ID NO 4  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: HELIX  
LOCATION: 1..33  
OTHER INFORMATION: Rao and Argos identification method, potential helix  
NAME/KEY: HELIX  
LOCATION: 4..20  
OTHER INFORMATION: Klein, Kanehisa and Delisi identification method, potential  
NAME/KEY: HELIX  
LOCATION: 4..24  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F  
OTHER INFORMATION: helix  
NAME/KEY: MYRISTATE  
LOCATION: 12..16  
OTHER INFORMATION: Prosite match  
NAME/KEY: HELIX  
LOCATION: 50..70  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F  
OTHER INFORMATION: helix  
NAME/KEY: CARBOHYD  
LOCATION: 57..59  
OTHER INFORMATION: Prosite match  
NAME/KEY: HELIX  
LOCATION: 76..96  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F  
OTHER INFORMATION: helix  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 78  
OTHER INFORMATION: potential Tyrosine kinase site, Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 84  
OTHER INFORMATION: potential caseine kinase II site, Prosite match  
NAME/KEY: SITE  
LOCATION: 94..115  
OTHER INFORMATION: potential Leucine zipper site, Prosite match  
NAME/KEY: MYRISTATE  
LOCATION: 119..123  
OTHER INFORMATION: potential site, Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 133  
OTHER INFORMATION: potential protein kinase C, Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 147  
OTHER INFORMATION: potential caseine kinase II site, Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 194  
OTHER INFORMATION: potential protein kinase C, Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 215  
OTHER INFORMATION: potential Tyrosine kinase site, Prosite match  
NAME/KEY: SUPPARTITION  
LOCATION: 221

OTHER INFORMATION: Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 233  
OTHER INFORMATION: potential cAMP and cGMP dependant protein kinase site,  
OTHER INFORMATION: match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 235  
OTHER INFORMATION: potential caseine kinase II site, Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 306  
OTHER INFORMATION: potential protein kinase C, Prosite match  
NAME/KEY: HELIX  
LOCATION: 310..330  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
OTHER INFORMATION: helix  
NAME/KEY: MYRISTATE  
LOCATION: 319..323  
OTHER INFORMATION: Prosite match  
NAME/KEY: MYRISTATE  
LOCATION: 323..327  
OTHER INFORMATION: Prosite match  
NAME/KEY: AMIDATION  
LOCATION: 329  
OTHER INFORMATION: Prosite match  
NAME/KEY: HELIX  
LOCATION: 333..353  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
OTHER INFORMATION: helix  
NAME/KEY: MYRISTATE  
LOCATION: 341..345  
OTHER INFORMATION: Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 350  
OTHER INFORMATION: potential protein kinase C, Prosite match  
US-09-853-526-4

Query Match 100.0%; Score 1871; DB 22; Length 353;  
Best Local Similarity 100.0%; Pred. No. 4,1e-184;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVVLGTAFTVYVLMGVWRLLSAFPAFYQALDRLCYQVQSWLFFFEYTG 60  
DB 1 MRYLLPSVVLGTAFTVYVLMGVWRLLSAFPAFYQALDRLCYQVQSWLFFFEYTG 60  
QY 61 VOILLYDLDPRKKNENIYYLANHSTVDIMYADILAIRONALGHVRYVLEKGLKMLPLYGC 120  
DB 61 VOILLYDLDPRKKNENIYYLANHSTVDIMYADILAIRONALGHVRYVLEKGLKMLPLYGC 120  
QY 121 YFAOHGCIYKRSKAKFENEKEMRNKLOSYVDAGTPMYLVIPPEGTRVNPEDTKVLSASQAF 180  
DB 121 YFAOHGCIYKRSKAKFENEKEMRNKLOSYVDAGTPMYLVIPPEGTRVNPEDTKVLSASQAF 180  
QY 181 AAORGGLAVLKHVLPRIKATHVAFDCMKNYLDAYDTVYVEGDDGGORRESPTMEFL 240  
DB 181 AAORGGLAVLKHVLPRIKATHVAFDCMKNYLDAYDTVYVEGDDGGORRESPTMEFL 240  
QY 241 CKECPKIHIDRIDDKDVEDEQEHMRRLHEREIKDKMLIEFYESPDERRKRPFGKS 300  
DB 241 CKECPKIHIDRIDDKDVEDEQEHMRRLHEREIKDKMLIEFYESPDERRKRPFGKS 300  
QY 301 VNSKLSIKTKLPSMLILSGLTAGMLMTDAGRKLYVNTWYIGTLGCLMVTIKA 353  
DB 301 VNSKLSIKTKLPSMLILSGLTAGMLMTDAGRKLYVNTWYIGTLGCLMVTIKA 353

RESULT 8  
US-09-901-484A-4  
; Sequence 4, Application US/09901484A  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Chumakov, Ilya

```

APPLICANT: Bougueleret, Lydie
TITLE OF INVENTION: Prostate Cancer Gene
FILE REFERENCE: GEN-TLIXC3D2
CURRENT APPLICATION NUMBER: US/09/901,484A
CURRENT FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: US 08/996,306
PRIOR FILING DATE: 1997-12-22
PRIOR APPLICATION NUMBER: US 60/099,658
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: US 09/218,207
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: US 09/338,907
PRIOR FILING DATE: 1999-06-23
PRIOR APPLICATION NUMBER: US 09/853,526
PRIOR FILING DATE: 2001-05-11
NUMBER OF SEQ ID NOS: 578
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 353
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: HELIX
LOCATION: (1)..(33)
OTHER INFORMATION: Rao and Argos identification method, potential helix
NAME/KEY: HELIX
LOCATION: (4)..(20)
OTHER INFORMATION: Klein, Kanehisa and Delisi identification method, potential helix
NAME/KEY: HELIX
LOCATION: (4)..(24)
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
OTHER INFORMATION: potential helix
NAME/KEY: LIPID
LOCATION: (12)..(16)
OTHER INFORMATION: MYRISTATE, Prosite match
NAME/KEY: HELIX
LOCATION: (50)..(70)
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
OTHER INFORMATION: potential helix
NAME/KEY: CARBOHYD
LOCATION: (57)..(59)
OTHER INFORMATION: Prosite match
NAME/KEY: HELIX
LOCATION: (76)..(96)
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
OTHER INFORMATION: potential helix
NAME/KEY: MOD_RES
LOCATION: (78)..(78)
OTHER INFORMATION: PHOSPHORYLATION, potential Tyrosine kinase site, Prosite match
NAME/KEY: MOD_RES
LOCATION: (84)..(84)
OTHER INFORMATION: PHOSPHORYLATION, potential caseine kinase II site, Prosite match
NAME/KEY: SITE
LOCATION: (94)..(115)
OTHER INFORMATION: Potential leucine zipper site, Prosite match
NAME/KEY: LIPID
LOCATION: (119)..(123)
OTHER INFORMATION: MYRISTATE, Prosite match
NAME/KEY: MOD_RES
LOCATION: (133)..(133)
OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match
NAME/KEY: MOD_RES
LOCATION: (147)..(147)
OTHER INFORMATION: PHOSPHORYLATION, potential caseine kinase II, Prosite match
NAME/KEY: MOD_RES
LOCATION: (194)..(194)
OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match
NAME/KEY: MOD_RES
LOCATION: (215)..(215)
OTHER INFORMATION: PHOSPHORYLATION, potential Tyrosine kinase site, Prosite match
NAME/KEY: MOD_RES
LOCATION: (221)..(221)
OTHER INFORMATION: SULFATATION, Prosite match

```

```

NAME/KEY: MOD_RES
LOCATION: (233)..(233)
OTHER INFORMATION: PHOSPHORYLATION, potential cAMP and cGMP dependant protein kin
OTHER INFORMATION: site, Prosite match
NAME/KEY: MOD_RES
LOCATION: (235)..(235)
OTHER INFORMATION: PHOSPHORYLATION, potential caseine kinase II site, Prosite mat
NAME/KEY: MOD_RES
LOCATION: (306)..(306)
OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match
NAME/KEY: HELIX
LOCATION: (310)..(330)
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
OTHER INFORMATION: potential helix
NAME/KEY: LIPID
LOCATION: (319)..(323)
OTHER INFORMATION: MYRISTATE, Prosite match
NAME/KEY: LIPID
LOCATION: (323)..(327)
OTHER INFORMATION: MYRISTATE, Prosite match
NAME/KEY: MOD_RES
LOCATION: (329)..(329)
OTHER INFORMATION: AMIDATION, Prosite match
NAME/KEY: HELIX
LOCATION: (333)..(353)
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
OTHER INFORMATION: potential helix
NAME/KEY: LIPID
LOCATION: (341)..(345)
OTHER INFORMATION: MYRISTATE, Prosite match
NAME/KEY: MOD_RES
LOCATION: (350)..(350)
OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match
US-09-901-484A-4

```

Query Match 100.0%; Score 1871; DB 23; Length 353;

Best Local Similarity 100.0%; Pred. No. 4,1e-184; Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MRYLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYVQSMVLEFFENYTG 60
DB 1 MRYLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYVQSMVLEFFENYTG 60
QY 61 VOILLGDLPRKNKENIITLANHOSYVDVIADILAIRONALGHRVYLKEGKMLPYGC 120
DB 61 VOILLGDLPRKNKENIITLANHOSYVDVIADILAIRONALGHRVYLKEGKMLPYGC 120
QY 121 YFQHGIGIVKRSKFNEMRKLQSYDAGTPVLYTFPESTRYNPOTKVLASQAF 180
DB 121 YFQHGIGIVKRSKFNEMRKLQSYDAGTPVLYTFPESTRYNPOTKVLASQAF 180
QY 121 YFQHGIGIVKRSKFNEMRKLQSYDAGTPVLYTFPESTRYNPOTKVLASQAF 180
DB 121 YFQHGIGIVKRSKFNEMRKLQSYDAGTPVLYTFPESTRYNPOTKVLASQAF 180
QY 181 AAGRGIAVLKHLVTPRIKATHVAFDCMKNYLAIYDVTVVYEGKDDGGRRSPYTEFL 240
DB 181 AAGRGIAVLKHLVTPRIKATHVAFDCMKNYLAIYDVTVVYEGKDDGGRRSPYTEFL 240
QY 241 CKCEPKIHIDRIDKDDVDEDEHMRMLHEFEIKDKMLIEFYSPPERKRRPGKS 300
DB 241 CKCEPKIHIDRIDKDDVDEDEHMRMLHEFEIKDKMLIEFYSPPERKRRPGKS 300
QY 301 VNSKLSIKTKLPMLILSGITAGMLMTDAGRKLYVTWVIYGTLLGLMTYTKA 353
DB 301 VNSKLSIKTKLPMLILSGITAGMLMTDAGRKLYVTWVIYGTLLGLMTYTKA 353

```

RESULT 9

US-09-915-582-59  
Sequence 59, Application US/09915582

GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 17 Human Secreted Proteins  
FILE REFERENCE: P5723P1  
CURRENT APPLICATION NUMBER: US/09/915,582

;; CURRENT FILING DATE: 2001-07-27  
;; PRIOR APPLICATION NUMBER: PCT/US01/01431  
;; PRIOR FILING DATE: 2001-01-17  
;; PRIOR APPLICATION NUMBER: 60/179,065  
;; PRIOR FILING DATE: 2000-01-31  
;; PRIOR APPLICATION NUMBER: 60/180,628  
;; PRIOR FILING DATE: 2000-02-04  
;; PRIOR APPLICATION NUMBER: 60/231,968  
;; PRIOR FILING DATE: 2000-09-12  
;; NUMBER OF SEQ ID NOS: 97  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO: 59  
;; LENGTH: 353  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
;; US-09-915-582-59

Query Match 100.0%; Score 1871; DB 23; Length 353;  
Best Local Similarity 100.0%; Pred. No. 4, 1e-184;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTPATYVLAAGVWRLSAFLPAFYQALDRLCYQSVLFFFEKVTG 60  
DB 1 MRYLLPSVLLGTPATYVLAAGVWRLSAFLPAFYQALDRLCYQSVLFFFEKVTG 60  
QY 61 VOILLYDGLPRKNENIYLANHSTVDIYADILAIQNALGHVRYLKGLKMLPLYGC 120  
DB 61 VOILLYDGLPRKNENIYLANHSTVDIYADILAIQNALGHVRYLKGLKMLPLYGC 120  
QY 121 YFAOHGIIYVRSKAKFENEKEMRNKLOSVDAGTPMYLVIPPEGRNYPEQTKVLSAQAF 180  
DB 121 YFAOHGIIYVRSKAKFENEKEMRNKLOSVDAGTPMYLVIPPEGRNYPEQTKVLSAQAF 180  
QY 181 AAOGLAVLKHVLPRIKATHVAFDCKKNYLDALYDVTYVEGKDGGQRRSEPTMTTEFL 240  
DB 181 AAOGLAVLKHVLPRIKATHVAFDCKKNYLDALYDVTYVEGKDGGQRRSEPTMTTEFL 240  
QY 241 CKECPKIHIDRIDKKDVEEEDHMRMWLHERFEIKDKMLIEFYESPDERRRKRPFGKS 300  
DB 241 CKECPKIHIDRIDKKDVEEEDHMRMWLHERFEIKDKMLIEFYESPDERRRKRPFGKS 300  
QY 301 VNSKLSIKTKLPMSLLISGLTAGMLMTDAGRKLVTNTWYIGTLGCLMTVITKA 353  
DB 301 VNSKLSIKTKLPMSLLISGLTAGMLMTDAGRKLVTNTWYIGTLGCLMTVITKA 353

RESULT 10  
US-09-915-582-79  
;; Sequence 79, Application US/09915582  
;; GENERAL INFORMATION:  
;; APPLICANT: Rosen et al.  
;; TITLE OF INVENTION: 17 Human Secreted Proteins  
;; FILE REFERENCE: PS723PI  
;; CURRENT APPLICATION NUMBER: US/09/915,582  
;; PRIOR FILING DATE: 2001-07-27  
;; PRIOR APPLICATION NUMBER: PCT/US01/01431  
;; PRIOR FILING DATE: 2001-01-17  
;; PRIOR APPLICATION NUMBER: 60/179,065  
;; PRIOR FILING DATE: 2000-01-31  
;; PRIOR APPLICATION NUMBER: 60/180,628  
;; PRIOR FILING DATE: 2000-02-04  
;; PRIOR APPLICATION NUMBER: 60/231,968  
;; PRIOR FILING DATE: 2000-09-12  
;; NUMBER OF SEQ ID NOS: 97  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO: 79  
;; LENGTH: 353  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
;; US-09-915-582-79

Query Match 100.0%; Score 1871; DB 23; Length 353;  
Best Local Similarity 100.0%; Pred. No. 4, 1e-184;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTPATYVLAAGVWRLSAFLPAFYQALDRLCYQSVLFFFEKVTG 60  
DB 1 MRYLLPSVLLGTPATYVLAAGVWRLSAFLPAFYQALDRLCYQSVLFFFEKVTG 60  
QY 61 VOILLYDGLPRKNENIYLANHSTVDIYADILAIQNALGHVRYLKGLKMLPLYGC 120  
DB 61 VOILLYDGLPRKNENIYLANHSTVDIYADILAIQNALGHVRYLKGLKMLPLYGC 120  
QY 121 YFAOHGIIYVRSKAKFENEKEMRNKLOSVDAGTPMYLVIPPEGRNYPEQTKVLSAQAF 180  
DB 121 YFAOHGIIYVRSKAKFENEKEMRNKLOSVDAGTPMYLVIPPEGRNYPEQTKVLSAQAF 180  
QY 181 AAOGLAVLKHVLPRIKATHVAFDCKKNYLDALYDVTYVEGKDGGQRRSEPTMTTEFL 240  
DB 181 AAOGLAVLKHVLPRIKATHVAFDCKKNYLDALYDVTYVEGKDGGQRRSEPTMTTEFL 240  
QY 241 CKECPKIHIDRIDKKDVEEEDHMRMWLHERFEIKDKMLIEFYESPDERRRKRPFGKS 300  
DB 241 CKECPKIHIDRIDKKDVEEEDHMRMWLHERFEIKDKMLIEFYESPDERRRKRPFGKS 300  
QY 301 VNSKLSIKTKLPMSLLISGLTAGMLMTDAGRKLVTNTWYIGTLGCLMTVITKA 353  
DB 301 VNSKLSIKTKLPMSLLISGLTAGMLMTDAGRKLVTNTWYIGTLGCLMTVITKA 353

RESULT 11  
US-60-099-658-4  
;; Sequence 4, Application US/60099658  
;; GENERAL INFORMATION:  
;; APPLICANT: Cohen, Daniel  
;; APPLICANT: Chumakov, Ilya  
;; APPLICANT: Blumenfeld, Marta  
;; APPLICANT: Bouguetelat, Lydie  
;; TITLE OF INVENTION: Prostate cancer gene  
;; NUMBER OF SEQUENCES: 99  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESS: Knobe, Martens, Olson & Bear  
;; STREET: 501 West Broadway  
;; CITY: San Diego  
;; STATE: California  
;; COUNTRY: USA  
;; ZIP: 92101-3505  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy Disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: Win95  
;; SOFTWARE: Word  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/60/099,658  
;; FILING DATE:  
;; CLASSIFICATION:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Israelson, Ned A.  
;; REGISTRATION NUMBER: 29,655  
;; REFERENCE/DOCKET NUMBER: GENSET.018APR  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (619) 235-0176  
;; TELEFAX: (619) 235-0176  
;; INFORMATION FOR SEQ ID NO: 4:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 353 amino acids  
;; TYPE: AMINO ACID  
;; STRANDEDNESS: SINGLE  
;; TOPOLOGY: LINEAR  
;; MOLECULE TYPE: PROTEIN  
;; ORIGINAL SOURCE:  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: potential Transmembrane helix

```
LOCATION: 1..33
FEATURE:
IDENTIFICATION METHOD: Rao and Argos method
NAME/KEY: potential Transmembrane helix
LOCATION: 4..20
IDENTIFICATION METHOD: Klein, Kanehisa and Delisi method
FEATURE:
NAME/KEY: potential Transmembrane helix
LOCATION: 4..24
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony
and Wall method
FEATURE:
NAME/KEY: potential N-myristoylation site
LOCATION: 12
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Transmembrane helix
LOCATION: 50..70
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony and Wall method
FEATURE:
NAME/KEY: potential N-glycosylation site
LOCATION: 57
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Transmembrane helix
LOCATION: 76..96
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony and Wall method
FEATURE:
NAME/KEY: potential Tyrosine kinase phosphorylation site
LOCATION: 78
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Casein kinase II phosphorylation site
LOCATION: 84
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential leucine zipper pattern
LOCATION: 94..115
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential N-myristoylation site
LOCATION: 119
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Protein kinase C phosphorylation site
LOCATION: 133
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Casein kinase II phosphorylation site
LOCATION: 147
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Protein kinase C phosphorylation site
LOCATION: 194
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Tyrosine kinase phosphorylation site
LOCATION: 215
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Tyrosine sulfatation site
LOCATION: 221
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential cAMP- and cGMP-dependent protein kinase phosphorylation site
LOCATION: 233
IDENTIFICATION METHOD: prosite match
NAME/KEY: potential Casein kinase II phosphorylation site
LOCATION: 235
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Protein kinase C phosphorylation site
```

```
LOCATION: 306
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Transmembrane helix
LOCATION: 310..330
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony and Wall method
FEATURE:
NAME/KEY: potential N-myristoylation site
LOCATION: 319
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential N-myristoylation site
LOCATION: 323
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Amidation site
LOCATION: 329
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Transmembrane helix
LOCATION: 333..353
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komarony and Wall method
FEATURE:
NAME/KEY: potential N-myristoylation site
LOCATION: 341
IDENTIFICATION METHOD: prosite match
FEATURE:
NAME/KEY: potential Protein kinase C phosphorylation site
LOCATION: 350
IDENTIFICATION METHOD: prosite match
US-60-099-658-4
```

```
Query Match          100.0%  Score 1871;  DB 26;  Length 353;
Best Local Similarity 100.0%  Pred. No. 4.1e-184;
Matches 353;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Oy 1 MRYLLPSVVLGTAFTYVLANGVWRLSAFLPARYQALDDRLCYVOSMVLFFPENTG 60
Db 1 MRYLLPSVVLGTAFTYVLANGVWRLSAFLPARYQALDDRLCYVOSMVLFFPENTG 60
Oy 61 VOILLYDLPKNKENIIVLANHSTVDWIVADIIAIRONALGHVRYVLEKGLKMLPLYGC 120
Db 61 VOILLYDLPKNKENIIVLANHSTVDWIVADIIAIRONALGHVRYVLEKGLKMLPLYGC 120
Oy 121 YFAOHGGIYVKSASFENKERNKLOSYVDAGTPMIVLIFPEGTRYNPEDQTKVLSASQAF 180
Db 121 YFAOHGGIYVKSASFENKERNKLOSYVDAGTPMIVLIFPEGTRYNPEDQTKVLSASQAF 180
Oy 181 AAQRLAVLKHVLPRIKATVAFPCMKNYDAIVDVTVVEGKDDGGRRESPTMTEFL 240
Db 181 AAQRLAVLKHVLPRIKATVAFPCMKNYDAIVDVTVVEGKDDGGRRESPTMTEFL 240
Oy 241 CKCPKRIHIDRIDKDVPEQEHMRRLHREFEIKDKMLIEFESDPERRKRPFGKS 300
Db 241 CKCPKRIHIDRIDKDVPEQEHMRRLHREFEIKDKMLIEFESDPERRKRPFGKS 300
Oy 301 VNSKSIKTKLPMSLILSGITRAGMLMTDAGRKLYNTWYIYIGLLDCLWVTIKA 353
Db 301 VNSKSIKTKLPMSLILSGITRAGMLMTDAGRKLYNTWYIYIGLLDCLWVTIKA 353
```

```
RESULT 12
US-09-436-919-1
Sequence 1, Application US/09436919A
GENERAL INFORMATION:
APPLICANT: Leung, David W
TITLE OF INVENTION: Human Lysophosphatidic Acid Acyltransferase-epsilon
FILE REFERENCE: 1801B
CURRENT APPLICATION NUMBER: US/09/436,919A
CURRENT FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn Ver. 2.0
```



Db 312 VNSKLSIKKTLPSMLILSGTAGMLMTDAGRKLVTWITYGTLGCLWVTIKA 364

RESULT 15  
US-09-901-484A-5  
; Sequence 5, Application US/09901484A  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Chumakov, Ilya  
; APPLICANT: Bouquelere, Lydie  
; TITLE OF INVENTION: Prostate Cancer Gene  
; FILE REFERENCE: GEN-T11XC3D2  
; CURRENT APPLICATION NUMBER: US/09/901,484A  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: US 08/996,306  
; PRIOR FILING DATE: 1997-12-22  
; PRIOR APPLICATION NUMBER: US 60/099,658  
; PRIOR FILING DATE: 1998-09-09  
; PRIOR APPLICATION NUMBER: US 09/218,207  
; PRIOR FILING DATE: 1998-12-22  
; PRIOR APPLICATION NUMBER: US 09/338,907  
; PRIOR FILING DATE: 1999-06-23  
; PRIOR APPLICATION NUMBER: US 09/853,526  
; PRIOR FILING DATE: 2001-05-11  
; NUMBER OF SEQ ID NOS: 578  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 364  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-901-484A-5

Query Match 100.0%; Score 1871; DB 23; Length 364;  
Best Local Similarity 100.0%; Pred. No. 4,2e-184;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVVLGTAPTYVLAMGVWRLSFLPARFYQALDDRLXCYYOSMVLFFPENNTG 60  
DB 12 MRYLLPSVVLGTAPTYVLAMGVWRLSFLPARFYQALDDRLXCYYOSMVLFFPENNTG 71  
QY 61 VOILLYGLDPKKNENIYLANHQSIVDMIVADILAIROMALGHVRVYLKEGLKMLPLYGC 120  
DB 72 VOILLYGLDPKKNENIYLANHQSIVDMIVADILAIROMALGHVRVYLKEGLKMLPLYGC 131  
QY 121 YFAOHGGIYVKRSKAFNEKEMRNKLQSYVDAGTPMYLVIFPEGTRYNPQOTKVLASQAF 180  
DB 132 YFAOHGGIYVKRSKAFNEKEMRNKLQSYVDAGTPMYLVIFPEGTRYNPQOTKVLASQAF 191  
QY 181 AAOBGLAVLKHYLTPRIKATHAFFDCMKNYLDAIVDTVVYEGKDDGQRRSPITTEFL 240  
DB 192 AAOBGLAVLKHYLTPRIKATHAFFDCMKNYLDAIVDTVVYEGKDDGQRRSPITTEFL 251  
QY 241 CKECPKIHIDIRIDKDVPEEOEHMRRLHERFEIKDKMLIEFYSPDPERRRKRPFGKS 300  
DB 252 CKECPKIHIDIRIDKDVPEEOEHMRRLHERFEIKDKMLIEFYSPDPERRRKRPFGKS 311  
QY 301 VNSKLSIKKTLPSMLILSGTAGMLMTDAGRKLVTWITYGTLGCLWVTIKA 353  
DB 312 VNSKLSIKKTLPSMLILSGTAGMLMTDAGRKLVTWITYGTLGCLWVTIKA 364

Search completed: August 28, 2002, 11:14:10  
Job time: 489 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:06:26 ; Search time 84.08 Seconds  
(without alignments)  
1016.222 Million cell updates/sec

Title: US-09-853-526-4

Perfect score: 1871  
Sequence: 1 MRYLPSVVLGTAPTYVLA.....YVNTWLYGTLGLMTYIKR 353

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747981 seqs, 242050750 residues

Total number of hits satisfying chosen parameters: 747981

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%

Listing first 45 summaries

Database :

Pending\_Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/1/paa/PCF\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1871	100.0	364	US-10-184-648-63	Sequence 63, Appl
2	1868	99.8	353	US-09-629-469A-13028	Sequence 13028, A
3	1427	76.3	269	US-10-074-045-47	Sequence 47, Appl
4	390.5	20.9	375	US-09-935-625-11061	Sequence 11061, A
5	390.5	20.9	375	US-09-935-625-14197	Sequence 14197, A
6	390.5	20.9	375	US-09-935-625-14201	Sequence 14201, A
7	390.5	20.9	375	US-09-935-625-14254	Sequence 14254, A
8	390.5	20.9	375	US-09-935-625-23009	Sequence 23009, A
9	390.5	20.9	375	US-09-935-625-23025	Sequence 23025, A
10	387.5	20.7	351	US-09-935-625-11062	Sequence 11062, A
11	387.5	20.7	351	US-09-935-625-14198	Sequence 14198, A
12	387.5	20.7	351	US-09-935-625-14202	Sequence 14202, A
13	387.5	20.7	351	US-09-935-625-14255	Sequence 14255, A
14	387.5	20.7	351	US-09-935-625-23010	Sequence 23010, A
15	387.5	20.7	351	US-09-935-625-23026	Sequence 23026, A
16	387.5	20.7	375	US-09-935-625-11471	Sequence 11471, A
17	384.5	20.6	351	US-09-935-625-11472	Sequence 11472, A
18	374.5	20.0	343	US-09-935-625-11063	Sequence 11063, A
19	374.5	20.0	343	US-09-935-625-11473	Sequence 11473, A
20	374.5	20.0	343	US-09-935-625-14199	Sequence 14199, A
21	374.5	20.0	343	US-09-935-625-14203	Sequence 14203, A
22	374.5	20.0	343	US-09-935-625-14256	Sequence 14256, A
23	374.5	20.0	343	US-09-935-625-23011	Sequence 23011, A
24	374.5	20.0	343	US-09-935-625-23027	Sequence 23027, A
25	371.5	19.9	400	US-60-391-781-1157	Sequence 1157, Ap
26	305.5	16.3	414	US-10-121-063-314	Sequence 314, App

27	305.5	16.3	414	US-10-063-502-102	Sequence 102, App
28	305.5	16.3	414	US-10-063-510-102	Sequence 102, App
29	305.5	16.3	414	US-10-063-512-102	Sequence 102, App
30	305.5	16.3	414	US-10-063-513-102	Sequence 102, App
31	305.5	16.3	414	US-10-063-514-102	Sequence 102, App
32	305.5	16.3	414	US-10-063-515-102	Sequence 102, App
33	305.5	16.3	414	US-10-063-516-102	Sequence 102, App
34	305.5	16.3	414	US-10-063-517-102	Sequence 102, App
35	305.5	16.3	414	US-10-063-518-102	Sequence 102, App
36	305.5	16.3	414	US-10-063-519-102	Sequence 102, App
37	305.5	16.3	414	US-10-063-520-102	Sequence 102, App
38	305.5	16.3	414	US-10-063-521-102	Sequence 102, App
39	305.5	16.3	414	US-10-063-522-102	Sequence 102, App
40	305.5	16.3	414	US-10-063-523-102	Sequence 102, App
41	305.5	16.3	414	US-10-063-524-102	Sequence 102, App
42	305.5	16.3	414	US-10-063-525-102	Sequence 102, App
43	305.5	16.3	414	US-10-063-526-102	Sequence 102, App
44	305.5	16.3	414	US-10-063-527-102	Sequence 102, App
45	305.5	16.3	414	US-10-063-528-102	Sequence 102, App

#### ALIGNMENTS

```
RESULT 1
US-10-184-648-63
: Sequence 63, Application US/10184648
:
: GENERAL INFORMATION:
: APPLICANT: Meyers, Rachel E.
: APPLICANT: Williamson, Mark
: APPLICANT: Tsai, Fong-Ying
: APPLICANT: Hunter, John J.
: APPLICANT: Macbeth, Kyle J.
: APPLICANT: Rudolph-Owen, Laura A.
: APPLICANT: Leiby, Kevin R.
: APPLICANT: Kapeller-Libermann, Rosana
: APPLICANT: Olandt, Peter J.
: TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE FAMILY MEMBERS AND USES THEREOF
: FILE REFERENCE: 10448-192001
: CURRENT APPLICATION NUMBER: US/10/184,648
: PRIOR FILING DATE: 2002-06-27
: PRIOR APPLICATION NUMBER: US 09/815,028
: PRIOR FILING DATE: 2001-03-22
: PRIOR APPLICATION NUMBER: PCT/US01/09358
: PRIOR FILING DATE: 2001-03-22
: PRIOR APPLICATION NUMBER: US 60/191,964
: PRIOR FILING DATE: 2000-03-24
: PRIOR APPLICATION NUMBER: US 09/801,220
: PRIOR FILING DATE: 2001-03-07
: PRIOR APPLICATION NUMBER: PCT/US01/07289
: PRIOR FILING DATE: 2001-03-07
: PRIOR APPLICATION NUMBER: US 60/187,456
: PRIOR FILING DATE: 2000-03-07
: PRIOR APPLICATION NUMBER: US 09/816,714
: PRIOR FILING DATE: 2001-03-23
: PRIOR APPLICATION NUMBER: PCT/US01/09468
: PRIOR FILING DATE: 2001-03-23
: PRIOR APPLICATION NUMBER: US 60/191,865
: PRIOR FILING DATE: 2000-03-24
: PRIOR APPLICATION NUMBER: US 09/844,948
: PRIOR FILING DATE: 2001-04-27
: PRIOR APPLICATION NUMBER: PCT/US01/13805
: PRIOR FILING DATE: 2001-04-27
: PRIOR APPLICATION NUMBER: US 60/200,604
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: US 09/861,164
: PRIOR FILING DATE: 2001-05-18
: PRIOR APPLICATION NUMBER: PCT/US01/16232
: PRIOR FILING DATE: 2001-05-18
: PRIOR APPLICATION NUMBER: US 60/205,408
: PRIOR FILING DATE: 2000-05-19
: PRIOR APPLICATION NUMBER: US 09/883,060
: PRIOR FILING DATE: 2001-06-15
```

```

; PRIOR APPLICATION NUMBER: PCT/US01/19138
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212, 079
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/962, 678
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: PCT/US01/29963
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/235, 044
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 09/973, 457
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: US 60/238, 849
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 10/072, 285
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: PCT/US02/03736
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/267, 494
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: US 09/817, 910
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: PCT/US01/09633
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/192, 092
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: US 09/842, 528
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: PCT/US01/40607
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 60/199, 500
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/882, 836
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19543
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/211, 730
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/882, 872
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19153
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212, 077
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-184-648-63
```

```

Query Match          100.0%; Score 1871; DB 6; Length 364;
Best Local Similarity 100.0%; Pred. No. 3.7e-167;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 MRYLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDRLRYCYQSMVLEFFENYTG 60
   |||||||
DB 12 MRYLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDRLRYCYQSMVLEFFENYTG 71
   |||||||
QY 61 VOILLYGLDLPKKNENITYLANHOSYVDWIVADILAIROMALGHVRVYLGKMLPLYGC 120
   |||||||
DB 72 VOILLYGLDLPKKNENITYLANHOSYVDWIVADILAIROMALGHVRVYLGKMLPLYGC 131
   |||||||
QY 121 YFAOHGCIYVKSAAKNEKEMRNKLOSYYVDAGTPTMYLVIFPEGTRYNPQOTVLSASQAF 180
   |||||||
DB 132 YFAOHGCIYVKSAAKNEKEMRNKLOSYYVDAGTPTMYLVIFPEGTRYNPQOTVLSASQAF 191
   |||||||
QY 181 AAORGLAVLKHVLTPIRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDGGORRESPTMTFEL 240
   |||||||
DB 192 AAORGLAVLKHVLTPIRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDGGORRESPTMTFEL 251
   |||||||
```

```

QY 241 CKECPKIHIDRIDKDDVPEOEHRMRWLHERFEIKDKMLIEFYSPDERKKRPPGKS 300
   |||||||
DB 252 CKECPKIHIDRIDKDDVPEOEHRMRWLHERFEIKDKMLIEFYSPDERKKRPPGKS 311
   |||||||
QY 301 VNSKLSIKKTPSMLILSGITAGMLMTDAGRKLYVTMTYTGILGCLAWTIRA 353
   |||||||
DB 312 VNSKLSIKKTPSMLILSGITAGMLMTDAGRKLYVTMTYTGILGCLAWTIRA 364
   |||||||
```

```

RESULT 2
US-09-629-469A-13028
; Sequence 13028, Application US/09629469A
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: SAITO, KAORU
; APPLICANT: YAMAMOTO, JUNICHI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: WAKAMATSU, AI
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: OTSUKI, TETSUJI
```

```

; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
; FILE REFERENCE: 084335/0123
; CURRENT APPLICATION NUMBER: US/09/629,469A
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: JP 1999-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 1999-300253
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/159, 590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183, 322
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 19025
; SOFTWARE: PatencIn Ver. 2.1
; SEQ ID NO 13028
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-629-469A-13028
```

```

Query Match          99.8%; Score 1868; DB 5; Length 353;
Best Local Similarity 99.7%; Pred. No. 6.9e-167;
Matches 352; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 MRYLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDRLRYCYQSMVLEFFENYTG 60
   |||||||
DB 1 MRYLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDRLRYCYQSMVLEFFENYTG 60
   |||||||
QY 61 VOILLYGLDLPKKNENITYLANHOSYVDWIVADILAIROMALGHVRVYLGKMLPLYGC 120
   |||||||
DB 61 VOILLYGLDLPKKNENITYLANHOSYVDWIVADILAIROMALGHVRVYLGKMLPLYGC 120
   |||||||
QY 121 YFAOHGCIYVKSAAKNEKEMRNKLOSYYVDAGTPTMYLVIFPEGTRYNPQOTVLSASQAF 180
   |||||||
DB 121 YFAOHGCIYVKSAAKNEKEMRNKLOSYYVDAGTPTMYLVIFPEGTRYNPQOTVLSASQAF 180
   |||||||
QY 181 AAORGLAVLKHVLTPIRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDGGORRESPTMTFEL 240
   |||||||
DB 181 AAORGLAVLKHVLTPIRIKATHVAFDCMKNYLDAIYDVTVVYEGKDDGGORRESPTMTFEL 240
   |||||||
QY 241 CKECPKIHIDRIDKDDVPEOEHRMRWLHERFEIKDKMLIEFYSPDERKKRPPGKS 300
   |||||||
```

Db 241 CKECPRIHHIDRIDKDVPEEOEHMRMLHEFEIKDKMLIEFYESPDPERRRKREPGRKS 300  
QY 301 VNSKLSIKTKLPMSLLISGLTAGMLMTDAGRKLIVNTWYIGTLGCLMTWITKA 353  
Db 301 VNSKLSIKTKLPMSLLISGLTAGMLMTDAGRKLIVNTWYIGTLGCLMTWITKA 353

RESULT 3  
US-10-074-045-47  
; Sequence 47, Application US/10074045  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT221C1  
; CURRENT APPLICATION NUMBER: US/10/074,045  
; CURRENT FILING DATE: 2002-02-14  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 74  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 47  
; LENGTH: 269  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-074-045-47

Query Match 76.3%; Score 1427; DB 6; Length 269;  
Best Local Similarity 100.0%; Pred. No. 1,2e-125;  
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 TVDMIVADILAIHQNALGHVRYLKEGLKWLPLGYCYFAOHGSIYKRSKAFNEKEKRNK 144  
Db 1 TVDMIVADILAIHQNALGHVRYLKEGLKWLPLGYCYFAOHGSIYKRSKAFNEKEKRNK 60  
QY 145 LOSYVADGTPMYLVEPEGTRYNPEQTKVLSASQAFPAOAGLAVLKHVLPRIKATHVAF 204  
Db 61 LOSYVADGTPMYLVEPEGTRYNPEQTKVLSASQAFPAOAGLAVLKHVLPRIKATHVAF 120  
QY 205 DCKKNLDAIDYTVYVEEGKDGQRRESPTWTEFLCKECPKIHIDRIDKDVPEEOE 264  
Db 121 DCKKNLDAIDYTVYVEEGKDGQRRESPTWTEFLCKECPKIHIDRIDKDVPEEOE 180  
QY 265 HMRMLHEFEIKDKMLIEFYESPDPERRRKREPGRKS VNSKLSIKTKLPMSLLISGLTAGM 324  
Db 181 HMRMLHEFEIKDKMLIEFYESPDPERRRKREPGRKS VNSKLSIKTKLPMSLLISGLTAGM 240  
QY 325 LMTDAGRKLIVNTWYIGTLGCLMTWITKA 353  
Db 241 LMTDAGRKLIVNTWYIGTLGCLMTWITKA 269

RESULT 4  
US-09-935-625-11061  
; Sequence 11061, Application US/09935625  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE  
; FILE REFERENCE: 2750-1481P  
; CURRENT APPLICATION NUMBER: US/09/935,625  
; CURRENT FILING DATE: 2001-08-24  
; NUMBER OF SEQ ID NOS: 33136  
; SEQ ID NO 11061  
; LENGTH: 375  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: peptide  
; LOCATION: 1..375  
; OTHER INFORMATION: Ceres Seq. ID no. 1394370  
US-09-935-625-11061

Query Match 20.9%; Score 390.5; DB 5; Length 375;  
Best Local Similarity 26.9%; Pred. No. 3.9e-28;  
Matches 101; Conservative 78; Mismatches 128; Indels 69; Gaps 9;

QY 1 MRYLPSVVLGTAFTYVLAWG-----VWRLSAFLPAPFYQALDRLCYVQSNVLEFF 55  
Db 18 LRGIICLMTLVSTAFPMILIFMGFLSAVLRLEFS-----IRYSRKCVSFFFGSWLALPFLF 73  
QY 56 ENYTGVOILLYDLPKRNKEIITLANHOSVDMIVADILAIHQNALGHVRYLKEGLKWL 115  
Db 74 EKINKTKVIFSGDKVPCEDVLLIANHREVDMMYFMDLALRKGQGINIKYVLSKSLMKL 133  
QY 116 PLGYCYFAOHGSIYKRSKAFNEKERNKLOSVDAGTPMYLVEPEGTRYNPEQTKVLS 175  
Db 134 PLGMAFHLEFIFPERKMEVDANLRQIVSSFKDPDRLALMLAFPEGTYTEAKQ--- 190  
QY 176 ASQAFPAOAGLAVLKHVLPRIKATHVAFDCKMKNYDAIDYTVYVEEGKDGQRRESPT 235  
Db 191 RSKKFAENGLPILNNVLPRTKGFVSCLOELCSIDAVYDVITGYKTR----- 239  
QY 236 MTEFLCKECP-----KIHIDRIDKDVPEEOEHMRMLHEFEIKDKMLIE 283  
Db 240 -----CPSFLDNYGIEPSEVHIIIRINLTQIPNDEKIDINAMLMNTFOLKDLND 291  
QY 284 FYESPDERRRKREPGRKS VNSKLSIKTKLPMSLLISGLT-----AGMLMTDAGRKLX 334  
Db 292 FYSN-----GHFPNDETEKEFWNTKKYLINCLAVIAFTTICTLITPSSWTF-----RIT 341  
QY 335 VNTWYIGTLGCLMTV 350  
Db 342 VS-----LACYVLT 350

RESULT 5  
US-09-935-625-14197  
; Sequence 14197, Application US/09935625  
; GENERAL INFORMATION:  
; APPLICANT: N. ALEXANDROV et al.  
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA  
; FILE REFERENCE: 2750-1481P  
; CURRENT APPLICATION NUMBER: US/09/935,625  
; CURRENT FILING DATE: 2001-08-24  
; NUMBER OF SEQ ID NOS: 33136  
; SEQ ID NO 14197  
; LENGTH: 375  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: peptide  
; LOCATION: 1..375  
; OTHER INFORMATION: Ceres Seq. ID no. 3441314  
US-09-935-625-14197

Query Match 20.9%; Score 390.5; DB 5; Length 375;  
Best Local Similarity 26.9%; Pred. No. 3.9e-28;  
Matches 101; Conservative 78; Mismatches 128; Indels 69; Gaps 9;

QY 1 MRYLPSVVLGTAFTYVLAWG-----VWRLSAFLPAPFYQALDRLCYVQSNVLEFF 55  
Db 18 LRGIICLMTLVSTAFPMILIFMGFLSAVLRLEFS-----IRYSRKCVSFFFGSWLALPFLF 73  
QY 56 ENYTGVOILLYDLPKRNKEIITLANHOSVDMIVADILAIHQNALGHVRYLKEGLKWL 115  
Db 74 EKINKTKVIFSGDKVPCEDVLLIANHREVDMMYFMDLALRKGQGINIKYVLSKSLMKL 133  
QY 116 PLGYCYFAOHGSIYKRSKAFNEKERNKLOSVDAGTPMYLVEPEGTRYNPEQTKVLS 175  
Db 134 PLGMAFHLEFIFPERKMEVDANLRQIVSSFKDPDRLALMLAFPEGTYTEAKQ--- 190  
QY 176 ASQAFPAOAGLAVLKHVLPRIKATHVAFDCKMKNYDAIDYTVYVEEGKDGQRRESPT 235  
Db 191 RSKKFAENGLPILNNVLPRTKGFVSCLOELCSIDAVYDVITGYKTR----- 239

```
Db 191 RSKFFAENGILPIINVLPRTEKGFVSCLOELSCSLDAVYDVITGYTR----- 239
QY 236 MTEFLCKECP-----KIHIDRIDKKDVEEEOEHMRWLHEREFIKDKMLIE 283
Db 240 -----CPSFLDNVYGIERSEVHHIRIRNLQIINQOEKDINAMLMNTFOLKDQLND 291
QY 284 FYESPDERRRRKRPFGKSVNSKLSIKTLPMSLISGLT-----AGMLMTDAGRKL 334
Db 292 FYSN-----GHFNEGTEKEFNTRKYLINCLAVIAFTTICTHILTFESSMIWF---RIV 341
QY 335 VNTWITGTLGCLMVT 350
Db 342 VS-----LACYLYT 350

RESULT 6
US-09-935-625-14201
; Sequence 14201: Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE
; OF INVENTION: MODULATING VARIOUS RESPONSES
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935, 625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 14201
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; NAME/KEY: peptide
; LOCATION: 1..375
; OTHER INFORMATION: Ceres Seq. ID no. 3441318
US-09-935-625-14201

Query Match 20.9%; Score 390.5; DB 5; Length 375;
Best Local Similarity 26.9%; Pred. No. 3.9e-28;
Matches 101; Conservative 78; Mismatches 128; Indels 69; Gaps 9;

QY 1 MRYLLPSVILGTAPTYVLAMG-----VWRLLSAFLPAREYQALDDRLXCVQSMVLFF 55
Db 18 LRGITCLMVLVSTAFPMMLIFWGLSAVVLRLFS---IRYSRKCVSFFFSWMLAMPPLF 73
QY 56 ENYTGVOILLYGDLPKKNENIYLANHSTVDWIVADILAIRONALGHVRYVLEKGLKWL 115
Db 74 EKIKTKTVIFSOGKVPEDRVLILANHRTEVDWMTFMDLALRKQIGINIKYVLSLMLK 133
QY 116 PLVGCYFAQHGGIYVKRSKAFNEKEMRNKLQSYVDAGTPMYLVIFPGSTRYINPEQTVLS 175
Db 134 PLFGMAFHLEFEIPIVERRWEDEANLRQIVSFKDPDALMLALFPEGTYTEAKCO--- 190
QY 176 ASQAFAROGIAVLKHYLTPRIKATHVAFDCMKNYLDAITDVYVYEGKDDGGRRRESPT 235
Db 191 RSKFFAENGILPIINVLPRTEKGFVSCLOELSCSLDAVYDVITGYTR----- 239
QY 236 MTEFLCKECP-----KIHIDRIDKKDVEEEOEHMRWLHEREFIKDKMLIE 283
Db 240 -----CPSFLDNVYGIERSEVHHIRIRNLQIINQOEKDINAMLMNTFOLKDQLND 291
QY 284 FYESPDERRRRKRPFGKSVNSKLSIKTLPMSLISGLT-----AGMLMTDAGRKL 334
Db 292 FYSN-----GHFNEGTEKEFNTRKYLINCLAVIAFTTICTHILTFESSMIWF---RIV 341
QY 335 VNTWITGTLGCLMVT 350
Db 342 VS-----LACYLYT 350

RESULT 7
US-09-935-625-14254
; Sequence 14254: Application US/09935625
```

```
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA
; OF INVENTION: MODULATING VARIOUS RESPONSES
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935, 625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 14254
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; NAME/KEY: peptide
; LOCATION: 1..375
; OTHER INFORMATION: Ceres Seq. ID no. 3443248
US-09-935-625-14254

Query Match 20.9%; Score 390.5; DB 5; Length 375;
Best Local Similarity 26.9%; Pred. No. 3.9e-28;
Matches 101; Conservative 78; Mismatches 128; Indels 69; Gaps 9;

QY 1 MRYLLPSVILGTAPTYVLAMG-----VWRLLSAFLPAREYQALDDRLXCVQSMVLFF 55
Db 18 LRGITCLMVLVSTAFPMMLIFWGLSAVVLRLFS---IRYSRKCVSFFFSWMLAMPPLF 73
QY 56 ENYTGVOILLYGDLPKKNENIYLANHSTVDWIVADILAIRONALGHVRYVLEKGLKWL 115
Db 74 EKIKTKTVIFSOGKVPEDRVLILANHRTEVDWMTFMDLALRKQIGINIKYVLSLMLK 133
QY 116 PLVGCYFAQHGGIYVKRSKAFNEKEMRNKLQSYVDAGTPMYLVIFPGSTRYINPEQTVLS 175
Db 134 PLFGMAFHLEFEIPIVERRWEDEANLRQIVSFKDPDALMLALFPEGTYTEAKCO--- 190
QY 176 ASQAFAROGIAVLKHYLTPRIKATHVAFDCMKNYLDAITDVYVYEGKDDGGRRRESPT 235
Db 191 RSKFFAENGILPIINVLPRTEKGFVSCLOELSCSLDAVYDVITGYTR----- 239
QY 236 MTEFLCKECP-----KIHIDRIDKKDVEEEOEHMRWLHEREFIKDKMLIE 283
Db 240 -----CPSFLDNVYGIERSEVHHIRIRNLQIINQOEKDINAMLMNTFOLKDQLND 291
QY 284 FYESPDERRRRKRPFGKSVNSKLSIKTLPMSLISGLT-----AGMLMTDAGRKL 334
Db 292 FYSN-----GHFNEGTEKEFNTRKYLINCLAVIAFTTICTHILTFESSMIWF---RIV 341
QY 335 VNTWITGTLGCLMVT 350
Db 342 VS-----LACYLYT 350

RESULT 8
US-09-935-625-23009
; Sequence 23009: Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA
; OF INVENTION: MODULATING VARIOUS RESPONSES
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935, 625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 23009
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; NAME/KEY: peptide
; LOCATION: 1..375
; OTHER INFORMATION: Ceres Seq. ID no. 3441318
US-09-935-625-23009
```



```
Sequence 14198, Application US/09935625
GENERAL INFORMATION:
APPLICANT: N. ALEXANDROV et al.
TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE
FILE REFERENCE: 2750-1481P
CURRENT APPLICATION NUMBER: US/09/935, 625
CURRENT FILING DATE: 2001-08-24
NUMBER OF SEQ ID NOS: 3136
SEQ ID NO 14198
LENGTH: 351
TYPE: PRP
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: peptide
LOCATION: 1..351
OTHER INFORMATION: Ceres Seq. ID no. 3441315
US-09-935-625-14198
```

```
Query Match 20.7%; Score 387.5; DB 5; Length 351;
Best Local Similarity 27.1%; Pred. No. 6.7e-28;
Matches 100; Conservative 75; Mismatches 125; Indels 69; Gaps 9;

QY 8 VLLGTAFTYVLWG-----VRLLSAFLPARFYQALDRLCYVOSMVLFFPENTGVQ 62
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 1 MVLVSTAFMMLIFWGLSAVLRIFS---IRSRKCVSFFFGSWLALWPFLEKINKTK 56

QY 63 ILLYGLPKKKNITLYLANHSTVDWIVADILAIRONALGHVRYVLEKGLPLYGCF 122
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 57 VIFSGDKVPCEDVLLIANHRTVDWMTFMDLARKGQIGNITYVLSKSLMKPLFGMAF 116

QY 123 AOHGIIYKRSKAFNEKEMRNKLOSVDAGTPMYLVIFPEGRYNEQTKVLSASQAF 182
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 117 HLEFIPVERKWEVDANLRQIVSFKPRDALMLALFPEGDTYTEAKQ---RSKKFAA 173

QY 183 QRCGLAVLKHVLTPIKATHTVAFCMKNYLDATIDVTVYVEGKDDGGORRESPMTEFLK 242
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 174 ENGLPILNNVLRPTKGFVSCLOELSCSDAVYDVITGYKTR----- 215

QY 243 ECP-----KIHIDRIDKKDVEDEGHRMRMLHERFETIKDKMLIEFYESPP 290
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 216 -CPSFDNYYGIPSEVHIIRINLTQIPNOKDINAMLMNTFQKDLNDFYSN--- 271

QY 291 ERRKRPFGKSVNSKLSIKKTLPSMLISGLT-----AGMLMTDAGKRLVNTWIYG 341
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 272 ---GHPNECTEKEFPMTKKYLINCLAVIAFTTICTHLPFSSIMWF---RIYVS----- 319

QY 342 TLGCLMVT 350
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 320 --LACVYLT 326
```

```
RESULT 12
US-09-935-625-14202
Sequence 14202, Application US/09935625
GENERAL INFORMATION:
APPLICANT: N. ALEXANDROV et al.
TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE
FILE REFERENCE: 2750-1481P
CURRENT APPLICATION NUMBER: US/09/935, 625
CURRENT FILING DATE: 2001-08-24
NUMBER OF SEQ ID NOS: 3136
SEQ ID NO 14202
LENGTH: 351
TYPE: PRP
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: peptide
LOCATION: 1..351
OTHER INFORMATION: Ceres Seq. ID no. 3441319
US-09-935-625-14202
```

```
Query Match 20.7%; Score 387.5; DB 5; Length 351;
Best Local Similarity 27.1%; Pred. No. 6.7e-28;
Matches 100; Conservative 75; Mismatches 125; Indels 69; Gaps 9;

QY 8 VLLGTAFTYVLWG-----VRLLSAFLPARFYQALDRLCYVOSMVLFFPENTGVQ 62
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 1 MVLVSTAFMMLIFWGLSAVLRIFS---IRSRKCVSFFFGSWLALWPFLEKINKTK 56

QY 63 ILLYGLPKKKNITLYLANHSTVDWIVADILAIRONALGHVRYVLEKGLPLYGCF 122
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 57 VIFSGDKVPCEDVLLIANHRTVDWMTFMDLARKGQIGNITYVLSKSLMKPLFGMAF 116

QY 123 AOHGIIYKRSKAFNEKEMRNKLOSVDAGTPMYLVIFPEGRYNEQTKVLSASQAF 182
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 117 HLEFIPVERKWEVDANLRQIVSFKPRDALMLALFPEGDTYTEAKQ---RSKKFAA 173

QY 183 QRCGLAVLKHVLTPIKATHTVAFCMKNYLDATIDVTVYVEGKDDGGORRESPMTEFLK 242
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 174 ENGLPILNNVLRPTKGFVSCLOELSCSDAVYDVITGYKTR----- 215

QY 243 ECP-----KIHIDRIDKKDVEDEGHRMRMLHERFETIKDKMLIEFYESPP 290
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 216 -CPSFDNYYGIPSEVHIIRINLTQIPNOKDINAMLMNTFQKDLNDFYSN--- 271

QY 291 ERRKRPFGKSVNSKLSIKKTLPSMLISGLT-----AGMLMTDAGKRLVNTWIYG 341
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 272 ---GHPNECTEKEFPMTKKYLINCLAVIAFTTICTHLPFSSIMWF---RIYVS----- 319

QY 342 TLGCLMVT 350
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 320 --LACVYLT 326
```

```
RESULT 13
US-09-935-625-14255
Sequence 14255, Application US/09935625
GENERAL INFORMATION:
APPLICANT: N. ALEXANDROV et al.
TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA
FILE REFERENCE: 2750-1481P
CURRENT APPLICATION NUMBER: US/09/935, 625
CURRENT FILING DATE: 2001-08-24
NUMBER OF SEQ ID NOS: 3136
SEQ ID NO 14255
LENGTH: 351
TYPE: PRP
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: peptide
LOCATION: 1..351
OTHER INFORMATION: Ceres Seq. ID no. 3443249
US-09-935-625-14255
```

```
Query Match 20.7%; Score 387.5; DB 5; Length 351;
Best Local Similarity 27.1%; Pred. No. 6.7e-28;
Matches 100; Conservative 75; Mismatches 125; Indels 69; Gaps 9;

QY 8 VLLGTAFTYVLWG-----VRLLSAFLPARFYQALDRLCYVOSMVLFFPENTGVQ 62
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 1 MVLVSTAFMMLIFWGLSAVLRIFS---IRSRKCVSFFFGSWLALWPFLEKINKTK 56

QY 63 ILLYGLPKKKNITLYLANHSTVDWIVADILAIRONALGHVRYVLEKGLPLYGCF 122
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 57 VIFSGDKVPCEDVLLIANHRTVDWMTFMDLARKGQIGNITYVLSKSLMKPLFGMAF 116

QY 123 AOHGIIYKRSKAFNEKEMRNKLOSVDAGTPMYLVIFPEGRYNEQTKVLSASQAF 182
   :||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 117 HLEFIPVERKWEVDANLRQIVSFKPRDALMLALFPEGDTYTEAKQ---RSKKFAA 173
```



---



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:04:46 ; Search time 39.13 Seconds  
(without alignments)  
866.842 Million cell updates/sec

Title: US-09-853-526-4

Perfect score: 1871

Sequence: 1 MRYLPSVVLGTAPTYVLA.....YVNWYIGTLGCLMTYIKR 353

Scoring table: BL0SUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 segs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR71:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	453	24.2	918	T34057	hypothetical prote
2	398.5	21.3	393	B96780	hypothetical prote
3	328	17.5	374	S52645	probable 1-acyl-g1
4	307	16.4	377	S60478	probable 1-acyl-sn
5	263	14.1	376	D96550	hypothetical prote
6	249.5	13.3	350	T40466	probable acetyltra
7	242.5	13.0	344	T31913	hypothetical prote
8	239	12.8	311	T07936	probable glycerol-
9	238.5	12.7	310	T06755	probable glycerol-
10	231.5	12.4	397	S45900	probable membrane
11	221.5	11.8	396	S54641	probable membrane
12	213.5	11.4	523	T25998	hypothetical prote
13	212	11.3	439	T22689	hypothetical prote
14	182	9.7	391	T15366	hypothetical prote
15	178	9.5	295	B83587	probable polynucle
16	154	8.2	310	S40808	polynucleotide ade
17	154	8.2	310	A91227	probable endonuc
18	154	8.2	310	G86074	probable endonuc
19	152.5	8.2	304	B83541	probable polynucle
20	139.5	7.5	285	T15252	hypothetical prote
21	135.5	7.2	302	A10950	probable acyltrans
22	131.5	7.0	363	T20608	hypothetical prote
23	130	6.9	294	D82371	probable polynucle
24	120.5	6.4	303	A48600	probable sn2-acylg
25	118	6.3	281	S60477	1-acylglycerol-3-p
26	114.5	6.1	247	G72223	hypothetical prote
27	114	6.1	266	G73863	probable 1-acylgly
28	112	6.0	258	G83101	probable acyltrans
29	111.5	6.0	488	OCBEHS	alkaline exonuclea

30	111	5.9	377	1	A44216	major envelope ant
31	109.5	5.9	488	2	T44030	alkaline exonuclea
32	109.5	5.9	488	2	T44215	alkaline exonuclea
33	109	5.8	240	2	D64089	1-acylglycerol-3-p
34	108	5.8	261	2	C82067	probable 1-acylgly
35	104	5.6	240	2	D64688	probable 1-acylgly
36	104	5.6	267	2	AC3499	1-acyl-sn-glycerol
37	103	5.5	243	2	AC0083	1-acylglycerol-3-p
38	103	5.5	266	2	B98283	hypothetical prote
39	103	5.5	266	2	AC3000	1-acyl-sn-glycerol
40	102	5.5	212	2	G85357	hypothetical prote
41	101	5.4	262	2	T22599	1-acylglycerol-3-p
42	101	5.4	323	2	C83940	sugar ABC transpor
43	99.5	5.3	237	2	B71827	probable 1-acylgly
44	99	5.3	225	2	B45582	probable 1-acylgly
45	99	5.3	243	2	B71706	probable 1-acylgly

## ALIGNMENTS

RESULT	1																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
--------	---	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Db 892 RTIGR-IYLLTIASSPLL-TAMLHIR 915

## RESULT 2

hypothetical protein F9E10.13 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001

C:Accession: B96780

R:Theologas, A.; Ecker, J.R.; Palm, C.J.; Pederspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Jensen, N.E.; Hughes, B.; Hultzer, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719

A:Accession: B96780

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1393 <STO>

A:Cross-references: GB:AE005173; NID:g6646762; PIDN:AAF21074.1; GSPDB:GN00141

C:Genetics:

A:Gene: F9E10.13

A:Map position: 1

Query Match 21.3%; Score 398.5; DB 2; Length 393;  
Best Local Similarity 29.8%; Pred. No. 1.5e-26;  
Matches 97; Conservative 68; Mismatches 117; Indels 43; Gaps 7;

1 MRYLPSVLLGTAAPYVLAGWRLSAFLPAFYQALDRLCYGVQSMV----- 52

23 LRGLMILVLSIAFMFL-----YFAP---TALGLRLSVQSKRVSLIGLW 70

53 -----FEFENVYGVQILLYDLPKNKENIYLANHSTVDMIVADILAIRONALGHVRYV 107

71 LALMPYLFETVNGTGVVFSDDIIPVEKRVLLIANHRTREVDMYLMIALKKGGLIKYV 130

108 LKGLKMLPRIGCYFAOHGIIYKRSKFNKEKRNKLOSVDAGTPMYLVIFPEGRYV 167

131 LKSLMLKLPFGMGFHVLEFIPVEKRVDEPVLQMLSSFKDQPEPLMLALPEEGDFT 190

168 PEQTKVLSAQAFARGLAVLKHLTPRIKATHVAFDCMKNYLDATYDVYVEGKDDG 227

191 EEKCK---RSQKFAEYGLALSIVLPTKRGVCLLEVLIHNSLDAYDLTIAYKPR--- 244

228 GORRESPTMTFLCKECP-KIHIIHDIRIDKRDVPEQEHMRMLHERFEIKDKMLIEFEY 286

245 -----CPSFDNVFGTDPSEVHIHVRVLLKEIPANEASSAMLMDSFKLKDLSDF-- 297

287 SPDPERRKRRPGKSVNSKLSIKTL 311

298 ----NAGKRPQRPEBELSVLKCI 318

## RESULT 3

probable 1-acyl-glycerol-3-phosphate acyltransferase - maize

C:Species: Zea mays (maize)

C:Date: 14-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 19-May-2000

C:Accession: S52645

R:Brown, A.P.; Coleman, J.; Tomney, A.M.; Watson, M.D.; Slabas, A.R.

Plant Mol. Biol. 26, 211-223, 1994

A:Title: Isolation and characterisation of a maize cDNA that complements a 1-acyl sn-gly-

her acyltransferase.

A:Reference number: S52645; MUID:95035993

A:Accession: S52645

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-374 <BRO>  
A:Cross-references: EMBL:Z29518; NID:g575959; PIDN:CAA82638.1; PID:g575960  
C:Superfamily: probable membrane protein YBR042c

Query Match 17.5%; Score 328; DB 2; Length 374;  
Best Local Similarity 29.8%; Pred. No. 1.8e-20;  
Matches 92; Conservative 64; Mismatches 121; Indels 32; Gaps 8;

18 VLMGVRMLISA-----FLPAR-----FYQALDRLCYGVQSMVLEFFENTGYQ 62

9 VLPGLGLPLISGLIVNAIQAVLFTTIPESKSFRRIRNRLAELMLQLVWVDMAGVK 68

63 ILLYGD-----LPKNENIYLANHSTVDMIVADILAIRONALGHVRYVLEKGLMPLY 118

69 VOLHADEYTRSMGKEHALIISNRSIDIMLIGMILAQSGCGLSTLAVAKSKSKFLPVI 128

119 G---CYFAOHGIIYKRSKFNKEKRNKLOSVDAGTPMYLVIFPEGRYVPEQTKYLSA 176

129 GWSMWFAYE--LFLERSMARKDEKTLKGLQRLKDFPRFWLALFVEGTRFPAK---LLA 183

177 SQAFPAQGRGLAVLKHLTPRIKATHVAFDCMKNYLDATYDVYVEGKDDGGRRRESPTM 236

184 AQEYASQGLPAPRNVLIPRTKGFSAVSTMRDPVATDTTVIVP-KDS-----PQPTM 237

237 TEFLCKECPKIHIIHDIRIDKRDVPEQEHMRMLHERFEIKDKMLIEFESPPERRKRP 296

238 LRILKQSSVIVHMKRHNAMSEMPKSDSDVSKKCKDIFVAKDALLKHLATGTFDEEIRP 297

297 PGKSVNSKL 305

298 IGRVKSLL 306

## RESULT 4

probable 1-acyl-sn-glycerol-3-phosphate acyltransferase - Limnanthes douglasii

C:Species: Limnanthes douglasii

C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 17-Mar-2000

C:Accession: S60478

R:Brown, A.P.; Brough, C.L.; Kroon, J.T.M.; Slabas, A.R.

Plant Mol. Biol. 29, 267-278, 1995

A:Title: Identification of a cDNA that encodes a 1-acyl-sn-glycerol-3-phosphate acylt

A:Reference number: S60477; MUID:96046746

A:Accession: S60478

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: mRNA

A:Residues: 1-377 <BRO>

A:Cross-references: EMBL:Z48730; NID:g1067137; PIDN:CAA88620.1; PID:g1067138

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1995

C:Superfamily: probable membrane protein YBR042c

Query Match 16.4%; Score 307; DB 2; Length 377;

Best Local Similarity 25.5%; Pred. No. 1.2e-18; Mismatches 131; Indels 50; Gaps 13;

26 LLSAFLPARFY---QALDRLCYGVQSMV-----LFFENVYGVQILLYDLPK--- 71

21 LVVNFIOAVFYVLYVPRISKTYRINTVLAELMLLELVWYIDMAGVKKQVLYDTSSFL 80

72 -NKENIYLANHSTVDMIVADILAIRONALGHVRYVLEKGLMPLYG--CYFAOHGI 128

81 MGKEHALIICNHRSDIDMLIGVLAQRCGLSSSIAMVKKSKFLPVYIGSMWFSEY--L 138

129 YVRSKAFNKEKRNKLOSVDAGTPMYLVIFPEGRYVPEQTKYLSA 187

139 FLERNNAKDKENTLKSGLQRLNDPPKPFWALFVEGTRF---TKAKLLAAQEYAAAGLP 194

188 VLKHLTPRIKATHVAFDCMKNYLDATYDVYVEGKDDGGORRESPTMTFLCKECPKI 247

195 VPRNVLPRTKGFSAVNSKSFVPAITDTVAIP-----KITTEOPTMLRLFRGSSVY 248



Db 122 -GAGMASSGSYIFLDRNFENDKPVLEIRIVKYSSGSEKKYOLLFAGETDKGERATRL- 178

QY 176 ASQAFARGLAVLKHVLTLPRIKATHVAFDCKM--NYLDAIYDVTVYEGKDDGGQRRES 233

Db 179 -SDAFKNGKLPRIYVHLHPTTFGFKLMELMKENIKIYVDITAYSG-----TIV 230

QY 234 PTMTFELCKECP-KIHIIHIDRIDKKDVPDEQEHMRMLHERFEIKDKMLIEFYSPD 289

Db 231 DTEAKLLAGNPPDKVHLDVKKYKRLDEIP-TGEGCEKMLTDLMATKERLKKFYEQEE 286

RESULT 8

T07936

probable glycerol-3-phosphate O-acyltransferase (EC 2.3.1.15) - rape

N:Alternate names: 1-acyl-sn-glycerol-3-phosphate acyltransferase

C:Species: Brassica napus (rape)

C>Date: 21-May-1999 #sequence\_revision 21-May-1999 #text\_change 20-Jun-2000

C:Accession: T07936

R:Brough, C.L.

submitted to the EMBL Data Library, June 1995

A:Accession number: 216230

A:Reference number: T07936

A:Status: preliminary; translated from GB/EMBL/DDBI

A:Molecule type: mRNA

A:Residues: 1-311 <BRO>

A:Cross-references: EMBL:249860; NID:g1149594; PIDN:CAA90019.1

A:Experimental source: cv. jet neuf; embryo

C:Superfamily: probable membrane protein YBR042C

C:Keywords: acyltransferase; coenzyme A

Query Match 12.8%; Score 239; DB 2; Length 311;

Best Local Similarity 29.8%; Pred. No. 6,8e-13;

Matches 59; Conservative 46; Mismatches 79; Indels 14; Gaps 5;

QY 86 VDMIVADIIILRONALCHVKYVLEKGLKMLPLYG--CYFAOHGCIYVKRSKFNKEKRN 143

Db 17 ISGLVNLID-RSGCLGSALAVMKSSKFLPVIGSMWFSEY--LFLERMMWAKDESTLKS 73

QY 144 KLSYVDAGTPMYLVIPPEGTQRYNPEQTKVLSASQAFARGLAVLKHVLTLPRIKATHVA 203

Db 74 GLQRLNDFPRFVNLALVEGTRFEAK---LKAQETVAASSELPVPRVNLIPRIKGVSA 130

QY 204 FDCMKNTLAIYDVTVYEGKDDGGQRRESPTMTFELCKECPKIHIIHIDRIDKKDVPDEQ 263

Db 131 VSNMRSPVPAIYDWTVAIP-----KTSPPPTMLRLFKGPSVYVHAIKCHSMKDEPSE 184

QY 264 EHMRLMLHERFEIKDKML 281

Db 185 DEIAQWCRDQFVTKDALL 202

RESULT 9

T06755

probable glycerol-3-phosphate O-acyltransferase (EC 2.3.1.15) - Arabidopsis thaliana

N:Alternate names: protein F15B8.160

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 19-May-2000

C:Accession: T06755

R:Queller, F.; Benes, V.; Rechmann, S.; Borkova, D.; Ansoerge, W.; Salanoubat, M.; Neues, submitted to the Protein Sequence Database, April 1999

A:Reference number: 215794

A:Accession: T06755

A:Molecule type: DNA

A:Residues: 1-310 <QUE>

A:Cross-references: EMBL:AL049660; GSPDB:GNO0061; ATSP:F15B8.160

A:Experimental source: cultivar Columbia; BAC clone F15B8

C:Genetics:

A:Gene: ATSP:F15B8.160

A:Map position: 3

A:Intons: 26/3; 46/3; 72/3; 125/3; 167/3; 198/3; 227/3

C:Superfamily: probable membrane protein YBR042C

C:Keywords: acyltransferase; coenzyme A

Query Match 12.7%; Score 238.5; DB 2; Length 310;

Best Local Similarity 24.8%; Pred. No. 7.5e-13;

Matches 67; Conservative 64; Mismatches 106; Indels 33; Gaps 8;

QY 97 RONALGHVRYVLEKGLMPLYG--CYFAOHGCIYVKRSKFNKEKRNKLSQYVDAGTP 154

Db 27 RSGCLGSALAVMKSSKFLPVIGSMWFSEY--LFLERMMWAKDESTLKSGLQRLSDPRP 84

QY 155 MYLVIPPEGTQRYNPEQTKVLSASQAFARGLAVLKHVLTLPRIKATHVADCKKNYIDAI 214

Db 85 FWLALFVEGTRFEAK---LKAQETVAASSELPVPRVNLIPRIKGVSAVSNMKSFPVPAI 141

QY 215 YDVTYVEGKDDGGQRRESPTMTFELCKECPKIHIIHIDRIDKKDVPDEQEHMRMLHERF 274

Db 142 YDMIVTIP-----KTSPPPTMLRLFKGPSVYVHAIKCHSMKMDLPESDAIAQWCRDQF 195

QY 275 EIKDKMLIEFYSPDPRRRRFPKGSVNSKLSIKTLPMLILSG-----LTAGML----- 325

Db 196 VAKDALDKHIAA-----DTPFQ--QEQNIGRPISKLAVLVSWACVLTGAIKFLHW 246

QY 326 --MTDAGRKLVTWITIGTLGCLMTYIKA 353

Db 247 AQLFSSWKGITISALGIGITLCMQILIRS 276

RESULT 10

S45900

probable membrane protein YBR042C - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein YBR0412

C:Species: Saccharomyces cerevisiae

C>Date: 26-Aug-1994 #sequence\_revision 09-Sep-1994 #text\_change 22-Oct-1999

C:Accession: S45900

R:Andre, B.; Cziepluch, C.; Hein, C.; Jauniaux, J.C.; Urrestiarazu, A.; Vissers, S. submitted to the Protein Sequence Database, August 1994

A:Reference number: S45903

A:Accession: S45900

A:Molecule type: DNA

A:Residues: 1-397 <AND>

A:Cross-references: EMBL:235911; NID:g536265; PIDN:CAA84984.1; PID:g536266; GSPDB:GNO

A:Experimental source: strain S288C

C:Genetics:

A:Gene: MIPS:YBR042C

A:Map position: 2R

C:Superfamily: probable membrane protein YBR042C

C:Keywords: transmembrane protein

F:12-37/Domain: transmembrane #status predicted <TM1>

F:55-77/Domain: transmembrane #status predicted <TM2>

F:134-150/Domain: transmembrane #status predicted <TM3>

F:372-390/Domain: transmembrane #status predicted <TM4>

Query Match 12.4%; Score 231.5; DB 2; Length 397;

Best Local Similarity 25.9%; Pred. No. 4.2e-12;

Matches 88; Conservative 55; Mismatches 136; Indels 61; Gaps 13;

QY 43 LYCYQSMVLEFFENYGVGILIXDLPKN-----KENIYLANHGSTVDWIVADLAI 96

Db 67 LHVAVPAAVAVHTTENSVPKGTFFLDLKKRRLISHLKSNSVAACNHOIYDWMFLMWLAY 126

QY 97 RONALGHVRYVLEKGLMPLYG--CYFAOHGCIYVKRSKFNKEKRNKL----- 145

Db 127 TSNUGANVFIILKKSILASITILFGMKNYFIFMSKRMADKTTLSNLSLAGLDSNARGAG 186

QY 146 -----QSYVDAGT-----PMVLVIFPEGTQRYNPEQTKVLSASQAFAPQ 183

Db 187 SLAKSPERITEEGESIMNDEVIDPKQIHWPYVNLIFPEGTNLSAD-TROKSAK--YAAK 243

QY 184 RGLAVLKHVLTLPRIKATHVAFDCKMNTLAIYDVTVYEG--KDDGGQRRESPTMTFELC 241

Db 244 IGKRPKNVLLPSTGLRYSLOKRLKPSIESLDYITIGSGVQKEGELIYG-LKSIPLF 302

```
OY 242 KECPR-1IHIDRIDKQVP-EEOEHMRRLHERFEIKMLIEFYES---DPERRRK 295
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 303 GKPRKLVDIRAFVDKDIPLDEDNFSEFLYIMSEKOLMERYSTGTFVSDPE---- 358

OY 296 FPKSVNSKLSIKK-TLPSMLISGLTAGMLMTDAGRKY 334
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 359 -TNHVTDSFKINRIELTEVLILPTLIILV-----YKLY 393

RESULT 11
S54641
probable membrane protein YDR018c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein D3246; hypothetical protein PAF396; hypothetical
C:Species: Saccharomyces cerevisiae
C>Date: 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 20-Jun-2000
C:Accession: S54641; S63425; S67831; S72116
R:Idem, K.; Brown, D.; Hamlyn, N.; Bowman, S.
submitted to the EMBL Data Library, May 1995
A:Reference number: S54638
A:Accession: S54641
A:Molecule type: DNA
A:Residues: 1-396 <DED>
A:Cross-references: EMBL:249770; NID:g940867; PIDN:CAA9843.1; PID:g940871
A:Experimental source: strain AB972
R:Idem, L.G.; Sander, C.; Prydz, H.
submitted to the EMBL Data Library, February 1996
A:Description: Sequencing and analysis of a 35.4 kb region on the left arm of chromosome
A:Reference number: S63416
A:Reference number: S63425
A:Accession: S63425
A:Molecule type: DNA
A:Residues: 1-396 <EID>
A:Cross-references: EMBL:X95966; NID:g1216215; PIDN:CAA65210.1; PID:g1216225
R:Prydz, H.; Eide, L.G.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67822
A:Accession: S67831
A:Molecule type: DNA
A:Residues: 1-396 <PRY>
A:Cross-references: EMBL:274314; NID:g1431443; PIDN:CAA9838.1; PID:g1431444; MIPS:YDR01
R:Experimental source: strain S286C
R:Eide, L.G.; Sander, C.; Prydz, H.
Yeast 12, 1085-1090, 1996
A>Title: Sequencing and analysis of a 35.4 kb region on the left arm of chromosome IV fr
A:Reference number: S72107; MUID:97051598
A:Accession: S72116
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-396 <EIM>
A:Cross-references: EMBL:X95966; NID:g1216215; PIDN:CAA65210.1; PID:g1216225
C>Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1996
C:Genetics:
A:Map position: 4R
A>Note: YDR018c
C:Superfamily: probable membrane protein YBR042c
C:Keywords: transmembrane protein
F:27-43/Domain: transmembrane #status predicted <TM1>
F:69-85/Domain: transmembrane #status predicted <TM2>
F:376-392/Domain: transmembrane #status predicted <TM3>

Query Match 11.8%; Score 221.5; DB 2; Length 396;
Best local Similarity 25.4%; Pred. No. 3, 1e-11;
Matches 78; Conservative 49; Mismatches 113; Indels 67; Gaps 10;

OY 73 KENIILANHOSTVDWIVADILAIKONALGHVRYLKEGKMLPLGYCYFAOHGGIYVKR 132
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 108 KDRALITIANHOMYADWIVLWMLSFVSNLGNVYIILKALQYIPLGFGMRNKFIFLSR 167

OY 133 SAKFNKEKMKNL-----QSYVDAGIPMT-LVIFPFGSTRYNPEQTK 172
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 168 NMOKDKRALTLNLSVMDLNARCKGPLTNKSKYSTINESIAAANLIMFPGNIMS--IK 224
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 173 VLSASQAFQAORGL--AVLKHVLTPIKATHVAFDCMKKNYLDIVDTVYVEGKDGGR 230
```

```
Db 225 TREKSEAFQRAHLNDVQLHLLPHSKGLKFAVEKLAPSLDAIVTIGY----- 275
OY 231 RESPTM-----TEFLCKEC-----PKIHIDRIDKQVP-EEOEHMRRLHERFEI 276
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 276 --SPALRTEYVGTFTLKKTFILMGVYPEKVDFFIRFERVNEIPLQDEVEFNNLLGVWKE 333
OY 277 KDKMLIEFESPPDERRRKPGKSVNSKLSIKKT-----LPSMLISGLTAGML 325
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 334 KDQLLEDYVMT-----GQFKSNAKNDNOSIVVTTGTFQHETLPRILISTYGFPAFLI 387

OY 326 MTDAGRK 332
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 388 LVFVWKK 394

RESULT 12
T25998
hypothetical protein ZK40.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T25998
R:Bradshaw, H.
submitted to the EMBL Data Library, July 1996
A:Description: The sequence of C. elegans cosmid ZK40.
A:Reference number: Z20121
A:Accession: T25998
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-523 <BRA>
A:Cross-references: EMBL:U64839; PIDN:AMB04844.1; GSPDB:GN00023; CESP:ZK40.1
A:Experimental source: strain Bristol N2; clone ZK40
C:Genetics:
A:Gene: CESP:ZK40.1
A:Map position: 5
A:Introns: 7/1; 38/3; 93/3; 126/3; 160/1; 301/3; 422/2; 472/3

Query Match 11.4%; Score 213.5; DB 2; Length 523;
Best local Similarity 22.5%; Pred. No. 2, 2e-10;
Matches 80; Conservative 62; Mismatches 148; Indels 65; Gaps 10;

OY 6 PSVVLGTAPTYVLAWGVMLLSAFLPARFYQALDRLYCVQSWLFFPENYTGVOILL 65
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 167 PLVVL-----FLKPRLMRQMDRLVGIWIMGALCNITFGANIRI 208

OY 66 YGDLPRKENIILANHOSTVDWIVADILAIKONALGHV-----RYLKEGKMLPL 117
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 209 KGDFINHEEPALITIMNRTRLDWL-----FVNALYKMDPWLCTTEKISLKGMLKYVP 262

OY 118 YGCFYAOHGGIYVKRSKAFKEKEMRKLQSYVDAGTPMYLVIFPEGTRYNPEQTYLSAS 177
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 263 AGWAMQAASYIFLDRSPDDTKLDNLNLYAETEKYQOLLPEEGTDCKPKATE--RS 319

OY 178 QAFAAORGAVLNLHVLTPRIKATHVAFDCMK--NYDAIYDVVYVEGKDGQORRESPT 235
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 320 RHSEKKGVLHYOVLHPRKTGEVHYIOMARRANNIKIYDVSIGF-----GDALYQSE 373

OY 236 MTEFLCKECPK-1IHIDRIDKQVPEEOEHMRRLHERFEIKMLIEFESPPDERRK 294
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 374 LDIFAHGVCPKVEFYQIKYPIEAIPTQDEALGOMVLNMRNKEKLRFEW--PRNVR 431

OY 295 RPFQGSVNSKLSIKTLPSMLISGLTAGMLMTDAGRKYLVNWIYGTLLGCLMW 349
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 432 QFPDTPDGVYEYLDNN-----TDRQKGLIGFWCFYTV--FMW 467

RESULT 13
T22689
hypothetical protein F55A11.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T22689
```



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:08:16 ; Search time 21.99 Seconds

(without alignments)  
621.555 Million cell updates/sec

Title: US-09-853-526-4

Perfect score: 1871

Sequence: 1 MRYLPSVVLGTAPRYVLA.....YVNTWIVGTLGLMTVTRKA 353

Scoring table:

BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1871	100.0	353	1 P1CE_HUMAN	Q9nqz2 homo sapien
2	1568.5	83.8	354	1 P1CE_MOUSE	Q9dl8e homo musculus
3	270	14.4	378	1 P1CD_RAT	Q9nrt51 rattus norv
4	268.5	14.4	378	1 P1CD_HUMAN	Q9nrt25 homo sapien
5	240	12.8	376	1 P1CC_HUMAN	Q9nrt27 homo sapien
6	231.5	12.4	397	1 YBP2_YEAST	P38226 saccharomyc
7	221.5	11.8	396	1 Y205_YEAST	Q92604 homo sapien
8	204.5	10.9	370	1 Y205_HUMAN	Q11087 caenorhabdi
9	182	9.7	391	1 Y773_CAEEL	P332129 escherichia
10	154	8.2	310	1 Y1NG_ECOLI	P333333 saccharomyc
11	120.5	6.4	303	1 P1SC_YEAST	Q42870 limnanthes
12	119	6.4	281	1 P1SC_LIMDO	Q42868 limnanthes
13	116	6.2	281	1 P1SC_LIMAL	P75479 mycoplasma
14	114	6.1	266	1 P1SC_MYCPN	P24447 human herpe
15	111.5	6.0	488	1 EXON_HSV6	P36316 fowlpox vir
16	111	5.9	377	1 VENV_FOXPV	P52448 human herpe
17	109.5	5.9	488	1 EXON_HSV6	P44448 haemophilus
18	109	5.8	240	1 P1SC_HAEIN	O25250 cocco nucif
19	104	5.6	240	1 P1SC_HELPI	Q93421 caenorhabdi
20	104	5.6	268	1 P1SC_COCNU	Q93441 caenorhabdi
21	101	5.4	302	1 P1C1_CAEEL	Q927n8 helicobacte
22	99.5	5.3	237	1 P1SC_HELPI	P26874 salmonella
23	99	5.3	245	1 P1SC_SALTY	Q09449 caenorhabdi
24	99	5.3	1551	1 YQ12_CAEEL	P26647 escherichia
25	97	5.2	245	1 P1SC_ECOLI	O10370 orgyia pseu
26	96	5.1	484	1 Y142_NPVOP	Q22667 caenorhabdi
27	94.5	5.1	282	1 P1C2_CAEEL	O59188 borrelia bu
28	93	5.0	280	1 P1SC_BORBU	P36617 schizosacch
29	90	4.8	892	1 RA16_SCHPO	O61542 mus musculu
30	88.5	4.7	446	1 MY64_MOUSE	O29052 archaeoglob
31	86	4.6	493	1 ALGO_ARCFU	P15275 pseudomonas
32	85.5	4.6	160	1 ALGO_PSEAR	P49902 homo sapien
33	85	4.5	561	1 SMTC_HUMAN	

34	84.5	4.5	255	1 P1SC_NEIMB	Q91247 neisseria m
35	84	4.5	1272	1 YZ28_METUA	O60287 methanococc
36	83.5	4.5	255	1 P1SC_METIC	O59601 neisseria g
37	83.5	4.5	255	1 P1SC_NEIMA	Q91041 neisseria m
38	83.5	4.5	268	1 P1SC_MYCGE	Q49402 mycoplasma
39	83.5	4.5	2386	1 RAD3_SCHPO	Q02099 schizosacch
40	83	4.4	318	1 YOK3_CAEEL	Q09287 caenorhabdi
41	83	4.4	778	1 YF05_METUA	O58900 methanococc
42	83	4.4	1241	1 TRK1_SACRA	P28569 saccharomyc
43	82.5	4.4	329	1 SUB1_SACRA	P06997 escherichia
44	82.5	4.4	998	1 PPO1_XENLA	P31669 xenopus lae
45	82	4.4	386	1 YMK7_YEAST	Q03760 saccharomyc

## ALIGNMENTS

RESULT ID	P1CE_HUMAN	STANDARD:	PRT:	353 AA.
AC	Q9NMQ2; Q9BQ64;			
DT	01-MAR-2002 (Rel. 41, Created)			
DT	01-MAR-2002 (Rel. 41, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	1-acyl-sn-glycerol-3-phosphate acyltransferase epsilon (EC 2.3.1.51)			
DE	(1-AGP acyltransferase 5) (1-AGPAT 5) (lysophosphatidic acid acyltransferase-epsilon) (LPAT-epsilon) (1-acylglycerol-3-phosphate O-acyltransferase 5).			
GN	AGPAT5.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Leung D.W.;			
RT	*Cloning and expression of LPAT-epsilon.*;			
RL	Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Cohen D., Chumakov I., Blumenfeld M., Bougueleret L.;			
RL	Patent number WO9332644, 01-JUL-1999.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Amgdala;			
RX	MEDLINE=21154917; PubMed=11230166;			
RA	Wiedmann S., Well B., Wellenreuther R., Gassenhuber J., Glassl S.,			
RA	Ansoorge W., Boecker M., Bloeker H., Bauersachs S., Blum H.,			
RA	Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,			
RA	Mewes H.-W., Othenwaelder B., Obermaier B., Tanpe J., Heubner D.,			
RT	Wambutt R., Korn B., Klein M., Poustka A.;			
RT	*Towards a catalog of human genes and proteins: sequencing and			
RT	analysis of 500 novel complete protein coding human cDNAs.*;			
RT	Genome Res. 11:422-435(2001).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Placenta;			
RA	Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,			
RA	Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,			
RA	Wagatsuna M., Hosokiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,			
RA	Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,			
RA	Matanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,			
RA	Yamamoto K., Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y.,			
RA	Nimomiya K., Iwayanagi T.;			
RT	*NEO human cDNA sequencing project.*;			
RT	Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.			
CC	-1- FUNCTION: CONVERTS LYOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC			
CC	ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION (BY			
CC	SIMILARITY).			
CC	-1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate -			
CC	CoA + 1,2-diacyl-sn-glycerol 3-phosphate.			
CC	-1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein (potential).			

CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
 CC ACYLTRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF375789; AAK54809.1; ALT\_INT.  
 DR EMBL: AL136587; CAB66522.1; ALT\_INT.  
 DR EMBL: AK002072; BAA92069.1; -  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KM Phospholipid biosyntheses; transferase; Acyltransferase;  
 KM Transmembrane.  
 FT TRANSMEM 7 29 POTENTIAL.  
 FT TRANSMEM 44 66 POTENTIAL.  
 FT TRANSMEM 334 351 POTENTIAL.  
 FT CONFLICT 145 145 L -> V (IN REF. 2).  
 SO SEQUENCE 353 AA; 40813 MW; A05B1FA246CE1B64 CRC64;

Query Match 100.0%; Score 1871; DB 1; Length 353;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-152;  
 Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRYLLPSVLLGTPPYVLAMGWRLLSAFLPARFYQALDRLCYQYQSVLFEFFNYTG 60  
 DB 1 MRYLLPSVLLGTPPYVLAMGWRLLSAFLPARFYQALDRLCYQYQSVLFEFFNYTG 60  
 QY 61 VOILLGDLPRKKNENIYYLANHSTYDVIADILAIROMLGHVRYLKEGLKMLPLGCG 120  
 DB 61 VOILLGDLPRKKNENIYYLANHSTYDVIADILAIROMLGHVRYLKEGLKMLPLGCG 120  
 QY 121 YFAOHGGIYVKKSAKFEKEMRNKLSYVDAGTPMYLVIFPESTRYNPQOTKVLASQAF 180  
 DB 121 YFAOHGGIYVKKSAKFEKEMRNKLSYVDAGTPMYLVIFPESTRYNPQOTKVLASQAF 180  
 QY 181 AAOGLAVLKHVLTLPRIKATHVAFDCKKNYLDIYDVTVYEEKDGQORRESPTMTEFL 240  
 DB 181 AAOGLAVLKHVLTLPRIKATHVAFDCKKNYLDIYDVTVYEEKDGQORRESPTMTEFL 240  
 QY 241 CKECPRKIHIDRIKDVPEEOEHMRWMLHERFEIKDKMLIEFYESPDERRRKRPFGKS 300  
 DB 241 CKECPRKIHIDRIKDVPEEOEHMRWMLHERFEIKDKMLIEFYESPDERRRKRPFGKS 300  
 QY 301 VNSKLSIKKTLPSMLISGLTAGMLTDAGRKLYVNTWITGTLGLCMTVITKA 353  
 DB 301 VNSKLSIKKTLPSMLISGLTAGMLTDAGRKLYVNTWITGTLGLCMTVITKA 353

RESULT 2  
 PILE\_MOUSE  
 ID PILE\_MOUSE STANDARD; PRT; 354 AA.  
 AC 09D1E8;  
 DT 01-MAR-2002 (Rel. 41, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE 1-acyl-sn-glycerol-3-phosphate acyltransferase epsilon (EC 2.3.1.51)  
 DE (1-AGP acyltransferase 5) (1-AGPAT 5) (Lysophosphatidic acid  
 DE acyltransferase-epsilon) (LPAAT-epsilon) (1-acylglycerol-3-phosphate  
 DE 0-acyltransferase 5).  
 GN AGPAT5.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Mus.  
 OX NCBI\_Taxid=10090;  
 RP [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Embryo;

RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,  
 RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Glass C., King B., Kochia H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schraml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustinich S., Hill D., Holman M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
 RA Hayashizaki Y.;  
 RA "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 CC -1- FUNCTION: CONVERTS LYSOEPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC  
 CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION (BY  
 CC SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate -  
 CC COA + 1,2-diacyl-sn-glycerol 3-phosphate.  
 CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).  
 CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
 CC ACYLTRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AK003649; BAB22915.1; -  
 DR EMBL: MG11915880; Agpat5.  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KW Phospholipid biosyntheses; transferase; Acyltransferase;  
 KM Transmembrane.  
 FT TRANSMEM 7 29 POTENTIAL.  
 FT TRANSMEM 44 66 POTENTIAL.  
 FT TRANSMEM 335 352 POTENTIAL.  
 SO SEQUENCE 354 AA; 40943 MW; 447E9F24B91E800 CRC64;

Query Match 83.8%; Score 1568.5; DB 1; Length 354;  
 Best Local Similarity 81.4%; Pred. No. 8.1e-127;  
 Matches 288; Conservative 39; Mismatches 26; Indels 1; Gaps 1;

QY 1 MRYLLPSVLLGTPPYVLAMGWRLLSAFLPARFYQALDRLCYQYQSVLFEFFNYTG 60  
 DB 1 MRYLLPSVLLGTPPYVLAMGWRLLSAFLPARFYQALDRLCYQYQSVLFEFFNYTG 60  
 QY 61 VOILLGDLPRKKNENIYYLANHSTYDVIADILAIROMLGHVRYLKEGLKMLPLGCG 120  
 DB 61 VOILLGDLPRKKNENIYYLANHSTYDVIADILAIROMLGHVRYLKEGLKMLPLGCG 120  
 QY 121 YFAOHGGIYVKKSAKFEKEMRNKLSYVDAGTPMYLVIFPESTRYNPQOTKVLASQAF 180  
 DB 121 YFAOHGGIYVKKSAKFEKEMRNKLSYVDAGTPMYLVIFPESTRYNPQOTKVLASQAF 180  
 QY 181 AAOGLAVLKHVLTLPRIKATHVAFDCKKNYLDIYDVTVYEEKDGQORRESPTMTEFL 239  
 DB 181 AAOGLAVLKHVLTLPRIKATHVAFDCKKNYLDIYDVTVYEEKDGQORRESPTMTEFL 240  
 QY 240 LCKECPKIHIDRIKDVPEEOEHMRWMLHERFEIKDKMLIEFYESPDERRRKRPFGK 299  
 DB 240 LCKECPKIHIDRIKDVPEEOEHMRWMLHERFEIKDKMLIEFYESPDERRRKRPFGK 299



[illegible]

```

Db 133 GMMWFTTE--VFCSSKMWODRRTVATSLQHLNDYEPKXYFFLIHCSTFTTEKHNEI--- 187
Oy 177 SOAFQAORGLAVLKHVLTFRIKATHVAFDCMKNNYDAIVTVVYEGKDDGGORRSPM 236
Db 188 SMOVAFRAKGLPRKHLHLPTKGFATVRSLSRVSAVVDCLINF-----RNENPPL 240
Oy 237 TEFLECKECPKIHIDRIKDDVPEQEHMRMLHERFEIKDKMLIEFY-----ESP-D 289
Db 241 LGLVANKKKYHADLYVRIFLEDIPEDDECSAMHLKLYQEKDAFQDEYVRTGTFPETPMV 300
Oy 290 PERRRFPKGSVN 302
Db 301 PPRR---PMTLVN 310

RESULT 5
PLCC_HUMAN STANDARD: PRT: 376 AA.
AC Q9NRZ7: Q9NRZ6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE 1-acyl-sn-glycerol-3-phosphate acyltransferase gamma (EC 2.3.1.51) (1-AGP acyltransferase 3) (1-AGPAT 3) (lysophosphatidic acid acyltransferase-gamma) (LPAT-gamma) (1-acylglycerol-3-phosphate O-acyltransferase 3).
DE acyltransferase-gamma).
GN AGPAT3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS GAMMA-1 AND GAMMA-2).
RA Leung D.W.;
RT "Structure and functions of lysophosphatidic acid acyltransferases.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal liver;
RA Nagamine K., Kudo J., Minoshima S., Kawasaki K., Hase T., Shimizu N.;
RT "Isolation of a novel gene encoding 1-acylglycerol-3-phosphate O-acyltransferase 3 (AGPAT3) from the human chromosome 21q22.3.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strussberg R.;
RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CONVERTS LYOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate -> COA + 1,2-diacyl-sn-glycerol 3-phosphate.
CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; GAMMA-1 (SHOWN HERE) AND GAMMA-2; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF156774; AAF80336.1; -
DR EMBL: AF156775; AAF80337.1; -
DR EMBL: AB040138; BAB18943.1; -

```

```

DR EMBL: BC011971; AAH11971.1; -
DR InterPro: IPR002123; Acyltransferase.
DR Pfam: PF01553; Acyltransferase; 1.
KW Phospholipid biosynthesis; Transferase; Acyltransferase;
KW Transmembrane; Alternative splicing.
FT TRANSMEM 15 39 POTENTIAL.
FT TRANSMEM 126 144 POTENTIAL.
FT TRANSMEM 308 329 POTENTIAL.
FT TRANSMEM 336 352 POTENTIAL.
FT VARSPIC 1 62 MISSING (IN ISOFORM GAMMA-2).
SQ SEQUENCE 376 AA; 43381 MW; C12CDBB7CC363852 CRC64;

Query Match 12.8%; Score 240; DB 1; Length 376;
Best Local Similarity 25.4%; Pred. No. 4e-13;
Matches 69; Conservative 63; Mismatches 116; Indels 24; Gaps 7;

Oy 34 RFYQALDDRLCYQSMVLFEEFENTGVQILYGLDPK-----NKENIILANHOSTVMI 89
Db 44 QLYRRLNCRLAYSLWSQVLVLEWMSCTECTLTQDAIVYERFKERAVIILHNEIDPL 103
Oy 90 VADIIAIPRONALGHVRYVLKEGLKMLPLYG--CYPAQHGGIYVKSAPKNEKEMNKLOS 147
Db 104 CGWTMCEKRGVLSGSKVLAKKELLYPLGLGWTWYFLEI--VECKRMEDRDVTYVGLLR 161
Oy 148 YVDAGTPMYLVIPFECTRYNPQOTKYLSAQAFAAORGLAVLKHVLTFRIKATHVAFDCM 207
Db 162 LSDPYEYMWFLLYCEGTRETFETKHHY---SMEVAAAKGLPVLKYLHLPTKGFTHAVKCL 218
Oy 208 KNYLDAIVDVTYVVEBKDDGGORRSPMTFELCKECPKIHIDRIKDDVPEQEHMR 267
Db 219 RGTVAAYVDVTNFRG-----NKNPSLGLIGLYKTKYADKCVRRFPLEDIPDEKEA 271
Oy 268 RMLHERFEIKDKMLIEFYESPDPERRKRRPK 299
Db 272 QMLHKIYQEKD-ALQEIY-----NOKGMPPGE 297

RESULT 6
YBP2_YEAST STANDARD: PRT: 397 AA.
AC P38226;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 45.5 kDa protein in FAY1-TCM62 intergenic region.
GN YBR042C OR YBR0412.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RA Andre B., Cziepluch C., Hein C., Jauniaux J.-C., Urrestarazu A., Vissers S.;
RT Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z35911; CAA84984.1; -
DR PIR: S45900; S45900.
DR SGD: S0000246; YBR042C.
DR InterPro: IPR002123; Acyltransferase.

```





DR EMBL: AE000461; AAC76860.1; -.  
 DR PIR: S40808; S40808.  
 DR Ecogene: EG11833; yihg.  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KM Hypothetical protein: Complete proteome.  
 SQ SEQUENCE 310 AA; 36289 MW; 9F8E3F52EB0B186E CRC64;

Query Match 8.2%; Score 154; DB 1; Length 310;  
 Best Local Similarity 20.6%; Pred. No. 7.2e-06;  
 Matches 64; Conservative 64; Mismatches 111; Indels 72; Gaps 14;

QY 1 MRYLLPSVLLGTAPFYVLAWSGWRLLSAFLPAREFOALDDLLCYQSMVFEFFENYTG 60  
 DB 43 VLLLPVPV-----IMRKVSF-----CDFPMKCCDELLAVLLHIN-PH 80  
 QY 61 VOILLYGLDPKKEN-IYLANHGSTVDWIVADILAIKONALGHV--RYVLEKGLKLP 116  
 DB 81 LQWEVHGLEGLSKKMYLLICNHRSMADIVLCVLFRR-----HIPMKYFLKQLAWVP 135  
 QY 117 LYG--C-----YFNOHGSIYKRSKAFKEMKRNKLSQSYVAGT-PMVLVIFPESTRN 167  
 DB 136 FVLGLACWSLIDMPKRYRSAYLLRPERGKDVETTRSCERFLAPPTIVNVEGSRPT 195  
 QY 168 PEOTFVLSASQAFAAORGIAVLKHLTPRIKATHVAFDCMKVYLDIYDVTVYVEGKDDG 227  
 DB 196 QEK-----HQQHTSTFQNLPPKAGIATMALNVLGQKDFKLAVTLCY--PDNN 242  
 QY 228 GORRESPTWTEFLCKECPRIHIDRIDKKDVEE-----QEHMRMLHERFEI 276  
 DB 243 RO-----PFEDMLSGKRLRIVH---VDLQPIADELHGDIYNDKSRKRFQQLSLMQE 294  
 QY 277 KDKMLIEFES 287  
 DB 295 KDLILSLMSS 305

RESULT 11  
 PLSC\_YEAST STANDARD: PRT; 303 AA.

AC P33333;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Probable 1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51)  
 DE (1-AGP acyltransferase) (1-AGPAT) (Lysophosphatidic acid  
 acyltransferase) (LPAAT).  
 GN SL1 OR YDL052C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 NC NCB1\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94012814; PubMed=8408076;  
 RA Nadiec M.M., Wells G.B., Lester R.L., Dickson R.C.;  
 RT "A suppressor gene that enables Saccharomyces cerevisiae to grow  
 without making sphingolipids encodes a protein that resembles an  
 Escherichia coli fatty acyltransferase";  
 RL J. Biol. Chem. 268:22156-22163(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX Urresterazu L.A., Andre B., Vissers S.;  
 RL Submitted (JUL-1996) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 182-303 FROM N.A.  
 RA Bloeker H., Brandt P.;  
 RL Submitted (JUL-1996) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: MAY BE AN ACYLTRANSFERASE WITH AN ALTERED SUBSTRATE  
 CC SPECIFICITY THAT ENABLES IT TO USE A C-26 COA IN PLACE OF THE  
 CC C-16 OR C-18 COAS USED BY THE WILD TYPE PROTEIN.  
 CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate =

COA + 1,2-diacyl-sn-glycerol 3-phosphate.  
 -1- PATHWAY: FIRST STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).  
 CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
 CC ACYLTRANSFERASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: L13282; AAA16514.1; -.  
 DR EMBL: Z74100; CAA98614.1; -.  
 DR PIR: A48600; A48600.  
 DR SGD: S0002210; SL1.  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KM Phospholipid biosynthesis; Transferase; Acyltransferase;  
 KW Transmembrane.  
 FT TRANSMEM 16 30 POTENTIAL.  
 FT VARIANT 44 44 Q -> L (IN SUPPRESSOR ALLELE SL1-1).  
 SQ SEQUENCE 303 AA; 33887 MW; 36CBBC2659655EB CRC64;

Query Match 6.4%; Score 120.5; DB 1; Length 303;  
 Best Local Similarity 29.7%; Pred. No. 0.0051;  
 Matches 44; Conservative 20; Mismatches 61; Indels 23; Gaps 6;

QY 45 CYQSMVLEFFENYTGVOILLYGLDLPKKENITYYLANHGSTVDWIVADILAIKONALGHV 104  
 DB 50 CYTHVWKML-----GLDKVYGEENLAKPIMINHOSTLIDIF-----LGR1 94  
 QY 105 -----RYVLEKGLKMLPLGYCYFAOAGIYKRSKAFKEMKRNKLSQSYVAGTPTMYTVI 159  
 DB 95 FPGGCTVTAKKSLKYVPFLGFMALSGTYFLDRSKROEALIDTLNKLLENVKK-NKRALMV 153  
 QY 160 PPEGTR-VNPEOTKVLASQAF-AAORG 185  
 DB 154 PPEGTRSYNSELTLMLPFKKGAFLAOG 181

RESULT 12  
 PLSC\_LIMDO STANDARD: PRT; 281 AA.

AC Q42870; Q40120;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE 1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51) (1-AGP  
 acyltransferase) (1-AGPAT) (Lysophosphatidic acid acyltransferase)  
 DE (LPAAT).  
 GN PLSC.  
 OS Limnantes douglasii (Douglas's meadowfoam).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Limnanthaceae; Limnantes.  
 NC NCB1\_TaxID=28973;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96028122; PubMed=7588719;  
 RA Hanke C., Wolter F.P., Coleman J., Peterek G., Frentzen M.;  
 RT "A plant acyltransferase involved in triacylglycerol biosynthesis  
 RT complements an Escherichia coli sn-1-acylglycerol-3-phosphate  
 RT acyltransferase mutant";  
 RL Eur. J. Biochem. 232:806-810(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96046746; PubMed=7579178;  
 RA Brown A.P., Brough C.L., Kroon J., Slabas A.R.;  
 RT "Identification of a cDNA that encodes a 1-acyl-sn-glycerol-3-

RT phosphate acyltransferase from Limnantes douglasii.";  
 RL Plant Mol. Biol. 29:267-278(1995).  
 CC -1- FUNCTION: CONVERTS LYOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC  
 CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION. THIS ENZYME  
 CC USES ERUCOYL-COA AS AN ACYL DONOR.  
 CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate -  
 CC CoA + 1,2-diacyl-sn-glycerol 3-phosphate.  
 CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
 CC ACYLTRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X83266; CAA58239.1; -  
 DR EMBL: Z46836; CAA6877.1; -  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KW Phospholipid biosynthesis; Transferase; Acyltransferase;  
 KM Transmembrane.  
 FT TRANSMEM 40 60 POTENTIAL.  
 FT TRANSMEM 71 91 POTENTIAL.  
 FT TRANSMEM 110 130 POTENTIAL.  
 FT CONFLICT 46 46 I -> V (IN REF. 2).  
 FT CONFLICT 188 188 R -> G (IN REF. 2).  
 FT CONFLICT 262 262 V -> I (IN REF. 2).  
 FT CONFLICT 281 281 N -> K (IN REF. 2).  
 SQ SEQUENCE 281 AA; 31716 MW; 9C8B0BD9E492EE2A CRC64;

Query Match 6.4%; Score 119; DB 1; Length 281;  
 Best Local Similarity 23.8%; Pred. No. 0.0062;  
 Matches 65; Conservative 40; Mismatches 92; Indels 76; Gaps 14;

QY 2 RYLLPSV-----VLGTAPTYV-----LAWGVRLSLFAPRFQAL 39  
 DB 13 KQLKPVAAATADDDKQGVNVLSCRFIECFRAIVLTAAMGL--IMVLLPMPYRIR 70  
 QY 40 DRLRY-CVYOSVLPFFENTYQVAILLYGDLPRKNENIYLANHOSVDIVADILAIRQ 98  
 DB 71 LGLYLGHIIIGGLIVIMY---GIPRIQGS-EHTKRAIYISNHSADIAFFVMWLA--- 122  
 QY 99 NALGHVRYVLKEGLKMLPLGYCFPAOHGITYVRSKAFENKRNKLQSYVDAGTP--MY 156  
 DB 123 -PIGTGVAKKEVI-WYPLLGLQYTLAHNIRIDRS--NPAAMIOSKKEAVRITERNLS 177  
 QY 157 LVTFPESTRYNEPQTVLASQAF---AAQRGLAVLKHVLTPIRIKATHVAF----- 204  
 DB 178 LIMEFPGTR--SNDGRLLPDKGCVHLALQSHLPVPMILT-----GTHLAWRKGTFRVRP 231  
 QY 205 -----DCMKNYLDIAYDVY 219  
 DB 232 VPIYKYLPIINTDWTVDKIDIDYVKMHHVYV 264  
 RESULT 13  
 PLSC\_LIML  
 ID PLSC\_LIML STANDARD; PRT; 281 AA.  
 AC 042868;  
 DT 15-DEC-1998 (rel. 37, Created)  
 DT 15-DEC-1998 (rel. 37, Last sequence update)  
 DT 15-DEC-1998 (rel. 37, Last annotation update)  
 DE 1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51) (1-AGP  
 DE acyltransferase) (1-AGPAT) (Lysophosphatidic acid acyltransferase)  
 DE (LPAAT).  
 OS Limnantes alba (White meadowfoam).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eucosids II; Brassicales; Limnathaceae; Limnantes.  
 OX NCBI\_Taxid=42439;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=ALBA;  
 RX MEDLINE=96106201; PubMed=8539298;  
 RA Lasserre M.W., Levering C.K., Davies H.M.D., Knutson D.S.;  
 RT "Lysophosphatidic acid acyltransferase from meadowfoam mediates  
 RT insertion of erucic acid at the sn-2 position of triacylglycerol in  
 RT transgenic rapeseed oil.";  
 RL Plant Physiol. 109:1389-1394(1995).  
 CC -1- FUNCTION: CONVERTS LYOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC  
 CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION. THIS ENZYME  
 CC USES ERUCOYL-COA AS AN ACYL DONOR.  
 CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate -  
 CC CoA + 1,2-diacyl-sn-glycerol 3-phosphate.  
 CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
 CC ACYLTRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U32968; AAC49185.1; -  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KW Phospholipid biosynthesis; Transferase; Acyltransferase;  
 KM Transmembrane.  
 FT TRANSMEM 40 60 POTENTIAL.  
 FT TRANSMEM 71 91 POTENTIAL.  
 FT TRANSMEM 110 130 POTENTIAL.  
 SQ SEQUENCE 281 AA; 31664 MW; 2725B81177A6067 CRC64;

Query Match 6.2%; Score 116; DB 1; Length 281;  
 Best Local Similarity 24.2%; Pred. No. 0.011;  
 Matches 59; Conservative 41; Mismatches 86; Indels 58; Gaps 13;

QY 9 VILGTAPTYVLANGVRLSAFLPAPFYQALDRLT-CVYOSVLPFFENTYQVAILLYG 67  
 DB 46 IYVLT-----VAMGL--IMVLLPMPYRIRLGNLYGIIIGGLIVIMY---GIPRIQGS 95  
 QY 68 DLPKNENIYLANHOSVDIVADILAIRQNALGHVRYVLKEGLKMLPLGYCFPAOHGG 127  
 DB 96 S-EHTKRAIYISNHSADIAFFVMWLA---PIGTGVAKKEVI-WYPLLGLQYTLAHN 149  
 QY 128 IYVRSKAFENKRNKLQSYVDAGTP--MYLVTFPESTRYNEPQTVLASQAF---AA 182  
 DB 150 IRLDRS--NPAAMIOSKKEAVRITERNLSLIMEFPGTRSG--DGRLLPDKGCVHLAL 204  
 QY 183 ORGLAVLKHVLTPIRIKATHVAF-----DCMKNYLDIAY 215  
 DB 205 QSHLPVPMILT-----GTHLAWRKGTFRVRPVITYKYLPINTDWTVDKIDIDYVKMH 260  
 QY 216 DVTV 219  
 DB 261 DIVV 264  
 RESULT 14  
 PLSC\_MYCPN  
 ID PLSC\_MYCPN STANDARD; PRT; 266 AA.  
 AC P75479;  
 DT 01-NOV-1997 (rel. 35, Created)  
 DT 16-OCT-2001 (rel. 40, Last sequence update)

```

DE Probable 1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51)
DE (1-AGP acyltransferase) (1-AGPAT) (lysophosphatidic acid
DE acyltransferase) (LPAT).
GN PLSC OR MPN29 OR MP537.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 29342 / M129;
RX MEDLINE=97105885; Pubmed=8948633;
RA Himmelfreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae."
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -1- FUNCTION: CONVERTS LYSOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC
CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION.
CC -1- CATALYTIC ACTIVITY: ACYL-CoA + 1-acyl-sn-glycerol 3-phosphate =
CC CoA + 1,2-diacyl-sn-glycerol 3-phosphate
CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.
CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE
CC ACYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000052; AAB96185.1; -
DR InterPro: IPR002123; Acyltransferase.
DR Pfam: PF01553; Acyltransferase; 1.
KW Phospholipid biosynthesis; Transferase; Acyltransferase;
KW Complete proteome.
SQ SEQUENCE 266 AA; 30414 MW; 79D933AD8203927F CRC64;

```

```

OC Betaherpesvirinae; Roseolovirus.
OX NCBI_TaxID=10370;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90080132; Pubmed=2152817;
RA Lawrence G.L., Chee M., Craxton M.A., Gompels U.A., Honess R.W.,
RA Barrell B.G.;
RT "Human herpesvirus 6 is closely related to human cytomegalovirus.";
RL J. Virol. 64:287-299(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95266321; Pubmed=7747482;
RA Gompels U.A., Nicholas J., Lawrence G., Jones M., Thomson B.J.,
RA Martin M.E., Efsthaliou S., Craxton M., Macaulay H.A.;
RT "The DNA sequence of human herpesvirus-6: structure, coding content,
RT and genome evolution."
RL Virology 209:29-51(1995).
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES ALKALINE EXONUCLEASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X83413; CA58362.1; -
DR EMBL: M68963; AA65578.1; -
DR PIR: F36769; QOBEHS.
DR InterPro: IPR001616; Herpes alk. exo.
DR Pfam: PF01771; Herpes_alk_exo.
DR PRINTS: PR00924; ALKEXNUCLASE.
KW Hydrolyase; Nuclease; Exonuclease.
SQ SEQUENCE 488 AA; 56644 MW; 0F38A10597366A5B CRC64;

```

Query Match 6.0%; Score 111.5; DB 1; Length 488;  
Best Local Similarity 20.6%; Pred. No. 0.056;  
Matches 71; Conservative 50; Mismatches 113; Indels 111; Gaps 16;

```

QY 23 WMRLSATLPARFOALDDRLCYQSMVLFEEFENVGVOLLGDDLPKKNITVLANH 82
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 134 LMHAL-----RKIVYATKLYLDITTRC---FLEYKQO--FRG-----EAVIYGAKH 176

QY 83 OSTVDWIVADILAIRONALGHVRYVLEKGLKWL-----PLYGCFYQHGSIY 129
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 177 ERVIRHLVATP-----YKREVKETGLLLDPSSGVGASLDACFGISFNEDDGLM 227

QY 130 VKRSA-----KFNKEMRNKLIQSYVD-----AGTPMYIVIFPEGTRNYPE 169
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 228 VKERKALFEIKFKYKYLDDKEDHVSSELKNPTKSFSDFTLSHPVPIEREREGKIPSS 287

QY 170 QTKVLSASQAAPAAORGALV--LKHVLPRI-----KATHVAFCKMKN----- 209
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 288 REYLMATYDFOYRPOKRLCTPTPALAPHIKOLLCLNFTQSKSYIVDFCKSDLEQKISV 347

QY 210 YLDAIYDVTVVYEGKDDGQRR-----ESPTMEFLCKECPKIHIDRIDKKDVP-- 260
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 348 FQKAVFTVNVFVNPK-----HRYFFQSLDQYVMTQFYIND-----HNNPEYIESTREYVSV 398

QY 261 -----EEOEHMRRLHREFEIKDKM---LIETYESPDPE 291
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 399 HVTVAFFRRRTREERSLTLVIDETEYIEEILPLALIVPVAPNDE 443

```

Search completed: August 28, 2002, 11:17:22  
Job time: 546 sec

( : \*



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:07:36 ; Search time 65.95 Seconds

(without alignments)  
925.962 Million cell updates/sec

Title: US-09-853-526-4

Perfect score: 1871

Sequence: 1 MRYLPSVVLGTAPTYVLA.....YVNTWYGTLLGLMTYKA 353

Scoring table:

BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_19:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1871	100.0	364	4 Q9BQ4	Q9BQ4 homo sapien
2	1868	99.8	353	4 Q9NUQ2	Q9NUQ2 homo sapien
3	1568.5	83.8	354	11 Q9DI18	Q9DI18 mus musculu
4	453	24.2	386	5 Q9SRI2	Q9SRI2 caenorhabdi
5	398.5	21.3	373	10 Q9SSH0	Q9SSH0 arabidopsis
6	398.5	21.3	393	10 Q9C9P8	Q9C9P8 arabidopsis
7	390.5	20.9	375	10 Q9LHNA	Q9LHNA arabidopsis
8	328	17.5	374	10 Q4I745	Q4I745 zea mays (m
9	309.5	16.4	390	10 Q9XFW4	Q9XFW4 brassica na
10	307	16.4	377	10 Q40119	Q40119 limnanthes
11	305.5	16.3	306	10 Q9SDN3	Q9SDN3 prunus dulc
12	272.5	14.6	386	5 Q9VY51	Q9VY51 drosophila
13	270	14.4	378	11 Q9Z4S1	Q9Z4S1 rattus norv
14	267	14.3	380	5 Q9VY49	Q9VY49 drosophila
15	263	14.1	376	10 Q9SYC8	Q9SYC8 arabidopsis
16	251	13.4	316	11 Q9DB84	Q9DB84 mus musculu

17	249.5	13.3	350	3 Q94361	Q94361 schizosacch
18	242.5	13.0	344	5 Q16526	Q16526 caenorhabdi
19	239	12.8	311	10 Q93317	Q93317 brassica na
20	238.5	12.7	310	10 Q9SVX9	Q9SVX9 arabidopsis
21	223	11.9	314	11 Q9D517	Q9D517 mus musculu
22	214	11.4	399	5 Q23087	Q23087 caenorhabdi
23	212	11.3	439	5 Q20800	Q20800 caenorhabdi
24	205.5	11.0	370	11 Q91YX5	Q91YX5 mus musculu
25	183.5	9.8	426	5 Q01882	Q01882 caenorhabdi
26	183.5	9.8	426	5 Q9SRO3	Q9SRO3 caenorhabdi
27	178	9.5	295	16 Q91657	Q91657 pseudomonas
28	152.5	8.2	304	16 Q915A6	Q915A6 pseudomonas
29	145	7.7	359	10 Q9FF57	Q9FF57 arabidopsis
30	131.5	7.0	363	5 Q19221	Q19221 caenorhabdi
31	130	6.9	294	16 Q9KVV5	Q9KVV5 vibrio chol
32	124	6.6	248	2 Q30849	Q30849 salmonella
33	116	6.2	281	10 Q9SDQ2	Q9SDQ2 limnanthes
34	114.5	6.1	247	16 Q9X219	Q9X219 thermotoga
35	113.5	6.1	385	5 Q9NFP1	Q9NFP1 leishmania
36	113	6.0	262	16 Q98G17	Q98G17 rhizobium 1
37	112	6.0	258	16 Q9CPE2	Q9CPE2 pasteurella
38	112	6.0	258	16 Q9HW50	Q9HW50 pseudomonas
39	109.5	5.9	235	2 Q9KX14	Q9KX14 rhodobacter
40	109.5	5.9	488	12 Q9WT03	Q9WT03 human herpe
41	108	5.8	261	16 Q9KPE3	Q9KPE3 vibrio chol
42	105.5	5.6	249	2 Q9EY25	Q9EY25 xanthomonas
43	105	5.6	260	16 Q9ZMG4	Q9ZMG4 rhizobium m
44	104	5.6	344	10 Q9LLY4	Q9LLY4 brassica na
45	102.5	5.5	291	3 Q9Y7C5	Q9Y7C5 emeritella

## ALIGNMENTS

RESULT 1

Q9BQ4 ID Q9BQ4 PRELIMINARY: PRT: 364 AA.

DT 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE HYPOTHETICAL 42.1 KDA PROTEIN (LYSOPHOSPHATIDIC ACID

DE ACYLTRANSFERASE-EPSILON) (EC 2.3.1.51).

GN DKFP761C222 OR LPAAT-E.

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP TISSUE=AMYGDALA;

RC MEDLINE=21154917; PubMed=11230166;

RX Wiemann S., Well B., Wellenreuther R., Gassenhuber J., Glass S.,

RA Ansgore W., Boecker M., Bloecher H., Bauersachs S., Blum H.,

RA Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Struck N.,

RA Mewes H.W., Oltmannseder B., Obermaier B., Tampe J., Heubner D.,

RA Mambrot R., Korn B., Klein M., Poustka A.;

RT "Towards a Catalog of Human Genes and Proteins: Sequencing and

RT Analysis of 500 Novel Complete Protein Coding Human cDNAs.";

RL Genome Res. 11:422-435(2001).

RN [2]

RP SEQUENCE FROM N.A.

RA Leung D.W.;

RT "Cloning and expression of LPAAT-epsilon.";

RT Submitted (May-2001) to the EMBL/Genbank/DBJ databases.

DR EMBL: AL136587; CAB66522.1; -

DR EMBL: AF375789; AAK54809.1; -

DR InterPro: IPR002123; Acyltransferase.

DR Pfam: PR01553; Acyltransferase: 1.

KW Hypothetical protein; transferase; Acyltransferase.

SQ SEQUENCE 364 AA: 42072 MW: 90A0F8FC7C8081 CRC64;

Query Match 100.0%; Score 1871; DB 4; Length 364;  
Best Local Similarity 100.0%; Pred. No. 4.2e-166;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRLLPSVLLGTAPTYVLANGVWRLLSAFRLPARFYQALDDRLCYCYQSVLFFFNNTG 60  
|||||  
DB 12 MRLLPSVLLGTAPTYVLANGVWRLLSAFRLPARFYQALDDRLCYCYQSVLFFFNNTG 71  
OY 61 VOILLYGDLPKKNENITYLANHOSVDMIVADILAIRQNALGHVRVYLEGKMLPLYGC 120  
|||||  
DB 72 VOILLYGDLPKKNENITYLANHOSVDMIVADILAIRQNALGHVRVYLEGKMLPLYGC 131  
OY 121 YFAOHGCIYVKSARKFNEKEMRNKLOSVDAGTPTMYLVFFPECTRYNPEOTKVLASQAF 180  
|||||  
DB 132 YFAOHGCIYVKSARKFNEKEMRNKLOSVDAGTPTMYLVFFPECTRYNPEOTKVLASQAF 191  
OY 181 AAOGLAVLKHVLTTPRIKATHVAFDCMKNYLDAIVDTVYVEGKDDGGORRESPTTEFL 240  
|||||  
DB 192 AAOGLAVLKHVLTTPRIKATHVAFDCMKNYLDAIVDTVYVEGKDDGGORRESPTTEFL 251  
OY 241 CECPKIHIDRIDKDDVPEOEHRMRWLHERFEIKDKMLIEFYESPDERRRKRPGRS 300  
|||||  
DB 252 CECPKIHIDRIDKDDVPEOEHRMRWLHERFEIKDKMLIEFYESPDERRRKRPGRS 311  
OY 301 VNSKLSIKKTLPSMLLSGLTAGMLTMDAGRKLYVTWTYIGTLGCLMTYTKA 353  
|||||  
DB 312 VNSKLSIKKTLPSMLLSGLTAGMLTMDAGRKLYVTWTYIGTLGCLMTYTKA 364

RESULT 2  
O9NUO2 PRELIMINARY; PRT: 353 AA.

AC O9NUO2;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)  
DE CDNA FLJ11210 FIS. CLONE PLACEL007954.  
OS Homo sapiens (Human).  
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PLACENTA;  
RA Iisogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
RA Wagaitsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
RA Matenabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y.,  
RA Niimura K., Iwayanagi T.;  
RT "NEO human cDNA sequencing project";  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK002072; BAA92069.1; -;  
DR InterPro: IPR002123; Acyltransferase.  
DR Pfam: PF01553; Acyltransferase; 1.  
SQ SEQUENCE 353 AA; 40799 MW; 0A8A14A2361F1B14 CRC64;

Query Match 99.8%; Score 1868; DB 4; Length 353;  
Best Local Similarity 99.7%; Pred. No. 7.7e-166;  
Matches 352; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRLLPSVLLGTAPTYVLANGVWRLLSAFRLPARFYQALDDRLCYCYQSVLFFFNNTG 60  
|||||  
DB 1 MRLLPSVLLGTAPTYVLANGVWRLLSAFRLPARFYQALDDRLCYCYQSVLFFFNNTG 60  
OY 61 VOILLYGDLPKKNENITYLANHOSVDMIVADILAIRQNALGHVRVYLEGKMLPLYGC 120  
|||||  
DB 61 VOILLYGDLPKKNENITYLANHOSVDMIVADILAIRQNALGHVRVYLEGKMLPLYGC 120  
OY 121 YFAOHGCIYVKSARKFNEKEMRNKLOSVDAGTPTMYLVFFPECTRYNPEOTKVLASQAF 180  
|||||

DB 121 YFAOHGCIYVKSARKFNEKEMRNKLOSVDAGTPTMYLVFFPECTRYNPEOTKVLASQAF 180  
OY 181 AAOGLAVLKHVLTTPRIKATHVAFDCMKNYLDAIVDTVYVEGKDDGGORRESPTTEFL 240  
|||||  
DB 181 AAOGLAVLKHVLTTPRIKATHVAFDCMKNYLDAIVDTVYVEGKDDGGORRESPTTEFL 240  
OY 241 CECPKIHIDRIDKDDVPEOEHRMRWLHERFEIKDKMLIEFYESPDERRRKRPGRS 300  
|||||  
DB 241 CECPKIHIDRIDKDDVPEOEHRMRWLHERFEIKDKMLIEFYESPDERRRKRPGRS 300  
OY 301 VNSKLSIKKTLPSMLLSGLTAGMLTMDAGRKLYVTWTYIGTLGCLMTYTKA 353  
|||||  
DB 301 VNSKLSIKKTLPSMLLSGLTAGMLTMDAGRKLYVTWTYIGTLGCLMTYTKA 353

RESULT 3  
O9D1E8 PRELIMINARY; PRT: 354 AA.

AC O9D1E8;  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)  
DE 1110013A05RIK PROTEIN.  
GN 1110013A05RIK  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=EMBRIO;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochava H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staubl F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamlya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Sasaki H., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyoko-Oka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,  
RA Wyszynski B., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection";  
RL Nature 409:685-690(2001).  
DR EMBL; AK003649; BAB22915.1; -;  
DR MGD; MGI:1915880; 1110013A05RIK.  
DR InterPro: IPR002123; Acyltransferase.  
DR Pfam: PF01553; Acyltransferase; 1.  
SQ SEQUENCE 354 AA; 40943 MW; 447EEF924B91E800 CRC64;

Query Match 83.8%; Score 1568.5; DB 11; Length 354;  
Best Local Similarity 81.4%; Pred. No. 6.6e-138;  
Matches 288; Conservative 39; Mismatches 26; Indels 1; Gaps 1;

OY 1 MRLLPSVLLGTAPTYVLANGVWRLLSAFRLPARFYQALDDRLCYCYQSVLFFFNNTG 60  
|||||  
DB 1 MRLLPSVLLGSAPTYVLANGLMTLWRLSLAMPARLXGRVDDRLCYCYQVMVFFFNNTG 60  
OY 61 VOILLYGDLPKKNENITYLANHOSVDMIVADILAIRQNALGHVRVYLEGKMLPLYGC 120  
|||||  
DB 61 VOILLYGDLPKKNENITYLANHOSVDMIVADILAIRQNALGHVRVYLEGKMLPLYGC 120  
OY 121 YFAOHGCIYVKSARKFNEKEMRNKLOSVDAGTPTMYLVFFPECTRYNPEOTKVLASQAF 180  
|||||

Db 121 YFAHGIGIYVKSAAKFNDEKMSKLSQSYVNAGMTVYIFPEGTRYNATYTKLISASQAF 180  
 QY 181 AAOBLAVLKHYLPRIKATHAFDCKMKNYLDAYDVTVVVEGKDDG-CGORESTPMTEF 239  
 Db 181 AAOBLAVLKHYLPRIKATHAFDCKMKNYLDAYDVTVVVEGKDDGSGKSYNPSMTEF 240  
 QY 240 LKEGCPKTHIHIDRDKDVEPEQEHMRMRMLHEREFEIKDKMLIEFESPDPERRRKFRFGK 299  
 Db 241 LCKGCPKTHIHIDRDKDVEPEQEHMRMRMLHEREFEIKDKMLIEFESPDPERRRKFRFGK 300  
 QY 300 SVNSKLSIKTKLPMSMLISGLTAGMLMTDAGRKLYVNTWIXYTGTLGCLMTVITKA 353  
 Db 301 SVHSRLSVKTKLPVYLIGSLTAVMLMTESGKLYMGTWLXGTLGCLMTVITKA 354

RESULT 4  
 ID Q95R12 PRELIMINARY; PRT; 386 AA.  
 AC Q95R12;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE HYPOTHEICAL 44.0 KDA PROTEIN.  
 GN F28B3.9.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 CX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA None;  
 RT "Genome sequence of the nematode C. elegans: a platform for  
 investigating biology. The C. elegans Sequencing Consortium.";  
 RL Science 282:2012-2018(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Gelsel C., Kramer J., Smith A.;  
 RT "The sequence of C. elegans cosmid F28B3.";  
 RL Submitted (May-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Waterston R.;  
 RT "Direct Submission.";  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF003136; AK93853.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 386 AA; 43978 MW; A1E3020C201452A6 CRC64;

Query Match 24.2%; Score 453; DB 5; Length 386;  
 Best Local Similarity 32.9%; Pred. No. 8.1e-34;  
 Matches 127; Conservative 72; Mismatches 139; Indels 48; Gaps 14;

QY 1 MYLLP-----SVVLGTAPTYVLANGVWRLLSAFLPARFYQALDRLCYVQSMVLFPEE 56  
 Db 12 LRPILPCSLSLSMVPFASCAIVIGV-----SWIVRHYAQLDMLMKYSYRRLCLFYEE 67  
 QY 57 NTGVOILLYGLDLPK--NK-----ENITYLANHSTVDNIVADIIAIRONALGH---VRY 107  
 Db 68 NISGFEITLHGNEEVYVNTKGFENAVNMISSHNSVNDMIIPMLARHGDDGNEQAFRYM 127  
 QY 108 LKEGKMLPLVGYCYFAOHGIGYVKSAAKFNDEKMSKLSQSYVNAGMTVYIFPEGTRYN 167  
 Db 128 VKNSHLVPMFGWYIFQGYIVRRGFEIGAPVLRQLKMLNESDPYVLLIFPEGTRNS 187  
 QY 168 PQOTVLSASQAFQAORGLAVLKHYLPRIKATHAFDCKMKNYLDAYDVTVVVEGKDD 226  
 Db 188 AKKHLLESSNRFLEKSGROPQWNVLCPRSGGLALDNLNST-LDAYDVTVVYQGMRD 246

QY 227 GG-----QRESPTWTEFLC--KECPKTHIHIDRDKDVEPEQEHMR 267  
 Db 247 LGLNLTTHYNLIITYRMAERKGLAPGMDPCCGQFQGLIHIDRIPIDVPRAKLETR 306  
 QY 268 RWLHEREFKDKMLIEFESPDPERRRKFRPKSVNSKLSIKTKLPMSMLIS-GLTAGMLM 326  
 Db 307 TWTLERFTKKEIRIDEFY-SEKPSGSAHP-----CVPISQTLPSLTLFSAALLAPFFS 359  
 QY 327 TDAGKRLVNVNTWIXYTGTLGCLMTVITK 352  
 Db 360 RTIGR-IYLLTASSPL-IAMLHIR 383

RESULT 5  
 ID Q9SSH0 PRELIMINARY; PRT; 373 AA.  
 AC Q9SSH0;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE F25A4.2 PROTEIN.  
 GN F25A4.2.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 CX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. COLUMBIA;  
 RA Vysotskaya V.S., Schwartz J.R., Yu G., Toriumi M., Lenz C., Liu S.,  
 RA Lee J.M., Li J., Gonzalez A., Liu A., Liu K., Vaysberg M., Sakano H.,  
 RA Chin C., Choi E., Chlou J., Altati H., Araujo R., Brooks S.,  
 RA Buehler E., Chao Q., Conn L., Conway A.B., Dunn P., Hansen N.,  
 RA Hwang B., Huizlar L., Khan S., Kim C., Palm C., Rowley D., Shinn P.,  
 RA Walker M., Davis R.W., Ecker J.R., Federspiel N.A., Theologis A.;  
 RT "Arabidopsis thaliana chromosome 1 BAC F25A4 sequence.";  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC008263; AAD55275.1; -  
 DR InterPro; IPR002123; Acyltransferase.  
 DR Pfam; PF01553; Acyltransferase; 1.  
 SQ SEQUENCE 373 AA; 42383 MW; 98642948D94709B6 CRC64;

Query Match 21.3%; Score 398.5; DB 10; Length 373;  
 Best Local Similarity 29.8%; Pred. No. 9.4e-29;  
 Matches 97; Conservative 68; Mismatches 117; Indels 43; Gaps 7;

QY 1 MYLLPVSVLGTAPTYVLANGVWRLLSAFLPARFYQALDRLCYVQSMVLFPEE----- 52  
 Db 23 LRGLMILVFLSTAFMFL-----YFAP---IALGLRLSVQSKRYVSLIFGLM 70  
 QY 53 -----FFENYTGVOILLYGLDLPKKNENIYLANHSTVDNIVADIIAIRONALGHVRY 107  
 Db 71 LALMPLRYLETGVNTGVFSGDIIPEKRVLLIANHRTVDWYIMNIALRKGCGLYRYV 130  
 QY 108 LKEGKMLPLVGYCYFAOHGIGYVKSAAKFNDEKMSKLSQSYVNAGMTVYIFPEGTRYN 167  
 Db 131 LKSSLMKLPPIFGWGFHLEFIPVERKREDEVLQMLSSFKDPEPMLALFPEGTF 190  
 QY 168 PQOTVLSASQAFQAORGLAVLKHYLPRIKATHAFDCKMKNYLDAYDVTVVVEGKDDG 227  
 Db 191 EKKCK-----RSQFAAEEVGLPALSNVLLPKTRGFCLEVLNNSLDAYVDTLTAQPR--- 244  
 QY 228 GQRESPTWTEFLKECP-KIHIIHIDRDKDVEPEQEHMRMRMLHEREFEIKDKMLIEYE 286  
 Db 245 -----CPSFMDVVFETDSEVHIHVRVLLKEIPANEAESSAMLDSEFKDKLSDP-- 297  
 QY 267 SPDPERRKRRPKGSVNSKLSIKTKL 311  
 Db 298 ---NAQGFPMQOREEELSLVKCI 318

RESULT 6  
Q9C9P8 PRELIMINARY: PRT: 393 AA.  
AC Q9C9P8: 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE PUTATIVE ACYL-COA:1-ACYLGLYCEROL-3-PHOSPHATE ACYLTRANSFERASE, 31588-29381.  
GN F9E10.13.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae: eucosids II: Brassicales: Brassicaceae: Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. COLUMBIA;  
RX MEDLINE=21016719; PubMed=11130712;  
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S., White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y., Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W., Chung M.K., Conn L., Conway A.B., Conway A.R., Cressy T.H., Dewar K., Dunn P., Eguu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y., Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L., Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E., Kim C.J., Koo H.L., Kremetska I., Kurtz D.B., Kwan A., Lam B., Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P., Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marzilli A., Miltsecher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I., Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D., Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M., Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D., Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M., Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;  
RA "Sequence and analysis of chromosome 1 of the plant Arabidopsis thaliana."  
RT thaliana.  
RL Nature 408:816-820(2000).  
DR EMBL: AC013258; AAG51931.1;  
DR InterPro: IPR002123; Acyltransferase.  
DR Pfam: PF01553; Acyltransferase: 1.  
KW Transferase: Acyltransferase.  
SQ SEQUENCE 393 AA; 44892 MW; FA5949D21E8E9A29 CRC64;

Query Match 21.3%; Score 398.5; DB 10; Length 393;  
Best Local Similarity 29.8%; Pred. No. 1e-28;  
Matches 97; Conservative 68; Mismatches 117; Indels 43; Gaps 7;  
QY 1 MYRLPSVLLGTAFTYVLAAGVWLLSAFLPARFYQALDDRLCYVQSNVLF----- 52  
DB 23 LRGLMILLVFLSTAFMFLL-----YFAP--IAALGLRLSVQGRKVVSLIFGLM 70  
QY 53 -----FFFNYYGVOILLXGDLPRKNENIYILANHOSTVDWIVADILATRONALGHVRY 107  
DB 71 LAIMPILFEVNGTIVVFGSDIIVPEKRVLLANHRTEDWMLMIALRKGLGKLYKYV 130  
QY 108 LKEGLMPLVLCGYFAQHGIIYKRSKFNKEMRNKLOSVDAGPMTLVIPPESTRN 167  
DB 131 LKSSLMKLLIFGFGHVLIEPIPERKREVDPEVLLQMLSSFKDPPRLMALPPEGTDT 190  
QY 168 PEOTKVLASQAFAAGRLAVLKHVLPRIKATHVAFDCKNLYDAIYDVTVVEGDKDG 227  
DB 191 EEKCK---RSQKFAAEVGLPALSNNVLPKTRGCGVCLLEVLSLDAVYDITAYKR--- 244  
QY 228 GGRRESPTMTPELCKECP-KIHIIHDIRDKDQVPEQEHMRRLHREFEIKMKLIEPE 286  
DB 245 -----CPSEFMDNVYFGTDPSEVHHVRRVLLKEIPANEAESSAMLMDSFKDKLSDF-- 297  
QY 287 SPDPERKRRPFGKSVNSKLSIKKTL 311  
DB 298 ----NAQGRFPNORPEEELSVLAKCI 318

RESULT 7  
Q9LHN4 PRELIMINARY: PRT: 375 AA.  
AC Q9LHN4: 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 15, Last annotation update)  
DE ACYL-COA:1-ACYLGLYCEROL-3-PHOSPHATE ACYLTRANSFERASE.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae: eucosids II: Brassicales: Brassicaceae: Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-COLUMBIA;  
RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;  
RL Submitted (MAY-2000) to the EMBL/Genbank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-COLUMBIA;  
RX MEDLINE=20363099; PubMed=10907853;  
RA Nakamura Y.;  
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II. Sequence features of the regions of 4,251,695 bp covered by ninety pl. TAC and BAC clones."  
RT DNA Res. 7:217-221(2000).  
DR EMBL: AP002039; BAB03094.1;  
DR InterPro: IPR002123; Acyltransferase.  
DR Pfam: PF01553; Acyltransferase: 1.  
KW Transferase: Acyltransferase.  
SQ SEQUENCE 375 AA; 43490 MW; 8F6F5DC313A3D704 CRC64;

Query Match 20.9%; Score 390.5; DB 10; Length 375;  
Best Local Similarity 26.9%; Pred. No. 5.3e-28;  
Matches 101; Conservative 78; Mismatches 128; Indels 69; Gaps 9;

QY 1 MYRLPSVLLGTAFTYVLAAGVWLLSAFLPARFYQALDDRLCYVQSNVLF----- 55  
DB 18 LRGLMILLVFLSTAFMFLLFVGLFSAVVLRLS-----IRYSKRVSPFGSLALMPFLF 73  
QY 56 ENYTGVOILLYDLPKKNENIYILANHOSTVDWIVADILATRONALGHVRYLAKGLMKL 115  
DB 74 EKINKKRVIFSGDKVCEEDRVLLANHRTEDWMLMIALRKGLGKLYKYVLSKSLKL 133  
QY 116 PLXGCFEAGHGGIYKRSKFNKEMRNKLOSVDAGPMTLVIPPESTRNPEQTKYLS 175  
DB 134 PLEGAHFHLPFTVPERKMEVEANLRQIVSSFKDPRDALMLALPPEGDIYEAQCO--- 190  
QY 176 ASQAFAGRLAVLKHVLPRIKATHVAFDCKNLYDAIYDVTVVEGDKDGGORRESPT 235  
DB 191 RSKKFAENGDLINNVLPKTRGFGVSCQELSCGLDANVYDITIGYKR----- 239  
QY 236 MTEFLCKECP-----KIHIIHDIRDKDQVPEQEHMRRLHREFEIKMKLIE 283  
DB 240 -----CPSELDNVYGIPESEVHHIRIRNLITOIPQEKDINAMLMTEFOLKQLIND 291  
QY 284 FYESPDERKRRPFGKSVNSKLSIKKTLPSMLISGLT-----AGMLTMDGRKLY 334  
DB 292 FYSN-----GHFPEGEKEKENTRKYLLINCLAVIAFTTICHTLFFSSMWF---RIV 341  
QY 335 VMTWITIGTLGLAWT 350  
DB 342 VS-----LACVYLT 350

RESULT 8  
Q41745 PRELIMINARY: PRT: 374 AA.  
AC Q41745: 01-NOV-1996 (TREMBLrel. 01, Created)  
DT

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE 1-ACYL-GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE (PUTATIVE).  
 OS Zea mays (Maize)  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
 OC Panicoideae; Andropogoneae; Zea.  
 OX NCBI\_TaxID=4577;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BLACK MEXICAN SWEET; TISSUE=ENDOSPERM;  
 RX MEDLINE=95035993; PubMed=7948871;  
 RA Brown A.P., Coleman J., Tommey A.M., Watson M.D., Slabas A.R.;  
 RT "Isolation and characterization of a maize cDNA that complements a 1-  
 acyl sn-glycerol-3-phosphate acyltransferase mutant of E.coli and  
 RT encodes a protein which has similarities to other acyltransferases.";  
 RL Plant Mol. Biol. 26:211-223(1994).  
 DR EMBL: Z29518; CAA82638.1; -;  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 DR PROSITE: PS00038; HELIX\_LOOP\_HELIX; UNKNOWN\_1.  
 KW Transferase; Acyltransferase.  
 SQ SEQUENCE 374 AA; 42571 MW; F1F5492CAF24F93 CRC64;

Query Match 17.5%; Score 328; DB 10; Length 374;  
 Best Local Similarity 29.8%; Pred. No. 3.5e-22;  
 Matches 92; Conservative 64; Mismatches 121; Indels 32; Gaps 8;

QY 18 VLAMGVMLLSA-----FLPAR-----FYQALDRLCYQSNVLPFFENYTGVO 62  
 DB 9 VLPGLFLSLGLYNAIOAVLFVTIRPFSKFYRINFLAELMLQLVWVDMWAGAK 68  
 QY 63 ILIYD-----LPKKNENIILYANHOSTVDWIVADILAIRONALGHVRYLAKGLKWLPLY 118  
 DB 69 VOLHADETRYRSMGKHEHLIISNHRSDIDWLIGMTIAQSGCLSTLAWMKSSKFLPYI 128  
 QY 119 G--CYFAOHGGIYVRSKAFNEKEMRNKLQSYVDAGTPMYLVIPEGPYRNEQIKVISA 176  
 DB 129 GMSMFAEY--LFLERSNAKDEKTLKWLQRLKDEPRPFWLALFVEGRTPEAK--LTA 183  
 QY 177 SOAFAAORGAVLKHLVTPRIKATHVAFDCMKNYLDIAYIVTVYVEGKDDGGORRESPTM 236  
 DB 184 AQEYASQGLPAPRNVLPRTKGFVSAYSIMRDFVPAITDTVIYP-KDS-----PQPTM 237  
 QY 237 TEFCKECKPIHIHIDRIDKVDPEQEHMRWLHERFEIKMKLIEFYESPDPERRKFE 296  
 DB 238 LRLKGSSVHIVRMKRRHAMSEMPKSDSDVSKWCKDIFVAKDALDLKHLATGTDFDEIRP 297  
 QY 297 PGKSVNSKL 305  
 DB 298 IGRPVKSL 306

RESULT 9  
 O9XFW4 PRELIMINARY; PRT; 390 AA.  
 AC O9XFW4;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE ACYL-COA:1-ACYLGLYCEROL-3-PHOSPHATE ACYLTRANSFERASE (EC  
 DE 2.3.1.51).  
 GN LPAAT.  
 OS Brassica napus (Rape).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Brassica.  
 OX NCBI\_TaxID=3708;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=DEVELOPING SILIQUES;

RA Graefin zu Muenster A., Wolter F.P., Frenzen M.;  
 RT "A cDNA encoding a microsomal 1-acylglycerol-3-phosphate  
 RT acyltransferase of Brassica napus L.";  
 RL Submitted (MAR-1997) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: Z95637; CAA09138.1; -;  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 DR PROSITE: PS00038; HELIX\_LOOP\_HELIX; UNKNOWN\_1.  
 KW Transferase; Acyltransferase.  
 SQ SEQUENCE 390 AA; 43771 MW; F1446EB30009C37 CRC64;

Query Match 16.5%; Score 309.5; DB 10; Length 390;  
 Best Local Similarity 26.5%; Pred. No. 2e-20;  
 Matches 83; Conservative 75; Mismatches 118; Indels 37; Gaps 9;

QY 58 YNGVOILYGD---LPKKNENIILYANHOSTVDWIVADILAIRONALGHVRYLAKGLK 113  
 DB 63 WAGVKIQVPADEFTFRMCKHEALVYCNHRSDIDWLIGMTIAQSGCLGSALAVMKSSK 122  
 QY 114 WPLUG--CYFAOHGGIYVRSKAFNEKEMRNKLQSYVDAGTPMYLVIPESTRNPEQT 171  
 DB 123 FLPIVGMWKFSEY--LFLERMAKDESTLQGLRLNDFPRFWLALFVEGRTPEAK- 179  
 QY 172 KYLSAQAFAAORGAVLKHLVTPRIKATHVAFDCMKNYLDIAYIVTVYVEGKDDGGORR 231  
 DB 180 --LKAQEYASSELPLVPARNVLIPRKGFVSAYSNRSEVPALYDTVAIP-----KTS 231  
 QY 232 ESPMTPEFLCKECPKIHIDRIDKVDPEQEHMRWLHERFEIKMKLIEFYESPDPPE 291  
 DB 232 PPTMLRLFKGGPSVHVHIIKCHSKMDLPEPDELAQWCRDGFVAKDALDLKHIAA---- 287  
 QY 292 RKRPRPKSVNSKLSIKTLPSMLISG---LTAGML-----MTDAGRLYNTWY 340  
 DB 288 --DTFPGQ---KEONIGRPKISLAVVSWACLLTIGAMFLHMSNLFSSWGIALSAFGL 342  
 QY 341 GTLGLCMWTYA 353  
 DB 343 GIITLCMQLINS 355

RESULT 10  
 O40119 PRELIMINARY; PRT; 377 AA.  
 ID O40119;  
 AC O40119;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE (PUTATIVE).  
 OS Limnanthes douglasii (Douglas's meadowfoam).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Limnathaceae; Limnanthes.  
 OX NCBI\_TaxID=28973;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=96046746; PubMed=7579178;  
 RA Brown A.P., Brough C.L., Kroon J., Slabas A.R.;  
 RT "Identification of a cDNA that encodes a 1-acyl-sn-glycerol-3-  
 RT phosphate acyltransferase from limnanthes douglasii.";  
 RL Plant Mol. Biol. 29:267-278(1995).  
 DR EMBL: Z48730; CAA86620.1; -;  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 DR PROSITE: PS00038; HELIX\_LOOP\_HELIX; UNKNOWN\_1.  
 KW Transferase; Acyltransferase.  
 SQ SEQUENCE 377 AA; 42780 MW; 0DECDFE25D39687C CRC64;

Query Match 16.4%; Score 307; DB 10; Length 377;  
 Best Local Similarity 25.5%; Pred. No. 3.3e-20;

Matches	91:	Conservative	85:	Mismatches	131:	Indels	50:	Gaps	13:	
Qy	26	LSLAFAPARF----	QALDRLCYQSNV-----	LEFFENYGOIILYGLPK---	71					
Db	21	LNVFICQAVFLVPRISKDTYRRINTLWVALLMELVWIDMWAGVQVQIYDTESFRL			80					
Qy	72	NKENIILAHOSVDMIVADILAIRONALGHVRVYLKELKMLPLYG--	CYFAOHGCI	128						
Db	81	MGEHALICHRSDIDMLIGWVLAORCGCLSSSIAMWKKSKFLPYIGMSMSEY--	L	138						
Qy	129	YVRSKAFENEKRNKLOSVDAGTSMYLYFEPTGRYNPQTFV--	LSASQAFAAORGIA	187						
Db	139	FLERNMADEWTEKSGQRLNDPFRFLALFVEGTRF----	TRAKLLAAQVYAASGLP	194						
Qy	188	VLVHVLPRITATVAVPDCMKNYIDAYDTVYVEGKDDGQRRSEPTMEFLCKECPKI		247						
Db	195	VPRVLLIPRTGFSVANSNMRSPFAPALDYLTVAIP-----	KTEQPTMLTLFRGKSSVY	248						
Qy	248	HIHIDRLDKVPEQOEHRMRWLEHFEIKDKMLIEFESDPERRRKRPCKSVNSKLSI		307						
Db	249	VHVLKRLMKLPKPTDDGVAAQWCKQFISKO-ALLDKHVAD----	TFSGLEVO--	DI 299						
Qy	308	KKTLPSMLILSG----	LTAGML-----	MTDAGRKLIVNVTGYCTGLGCMVTYKA	353					
Db	300	GRPKSLVYVYVSWMCLCLGLVLRFLQNSALLSSMKMMITTFVLCIYVLMHILIRS		356						
RESULT	11									
Q9SDN3		PRELIMINARY;	PRT;	306 AA.						
AC	Q9SDN3									
DT	01-MAY-2000 (TremBrel. 13, Created)									
DT	01-MAY-2000 (TremBrel. 13, Last sequence update)									
DT	01-DEC-2001 (TremBrel. 19, Last annotation update)									
DE	1-ACYL-SN-GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE.									
DE	Prunus dulcis (Almond) (Prunus amygdalus).									
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; Eucosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.									
OC	NCBI_Taxid=3755;									
RN	[1]									
RP	SEQUENCE FROM N.A.									
RT	Compalans A., Pages M., Messegue R.; "Identification of differentially expressed genes during dehydration in almond (Prunus amygdalus) using the cDNA-AFLP technique."; RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.									
DR	EMBL: AF213937; AAF20003.1;									
DR	Interpro: IPR002123; Acyltransferase.									
DR	Interpro: IPR003015; HLH_Myc.									
DR	Pfam: PF01553; Acyltransferase; 1.									
DR	PROSITE: PS00038; Helix_LOOP_Helix; UNKNOWN_1.									
DR	Transferrase; Acyltransferase.									
SO	SEQUENCE 306 AA; 33993 MW; CBD33449625908D CRC64;									
Query Match	16.38;	Score 305.5;	DB 10;	Length 306;						
Best Local Similarity	31.9%;	Pred. No. 3.3e-20;								
Matches	80;	Conservative 55;	Mismatches 101;	Indels 15;	Gaps 7;					
Qy	73	KENIYLANHOSVDMIVADILAIRONALGHVRVYLKELKMLPLYG--	CYFAOHGCIYV	130						
Db	3	KEHVLVLSNHSDDIDMLIGWVLAORSGCLSSLAVMKSKSFPLVIGMSMSEY--	LFL	60						
Qy	131	KRSKAFENEKRNKLOSVDAGTSMYLYFEPTGRYNPQTFV--	LSASQAFAAAGLAVLK	190						
Db	61	ERSMAKDEGTAKSGVQRLNDPFRFLALFVEGTRF----	QAKLL-AAQVYAASGLP	117						
Qy	191	VHVLPRITATVAVPDCMKNYIDAYDTVYVEGKDDGQRRSEPTMEFLCKECPKI	HIH	250						
Db	118	NVLIPRTGFSVANSNMRSPFAPALDYLTVAIP-----	KSSPAPTMLRLEFGREPSVHVH	171						
Qy	251	IDRLDKVPEQOEHRMRWLEHFEIKDKMLIEFESDPERRRKRPCKSVNSKLSIKK		309						
Db	172	IKRHVMNDLPEDEAVAQWCKDIFVAKDALLDKHNTYEQTFGQDQKLVATGRPLKS-LLVYT		23						
Qy	310	TLPSSLILSGL	320							
Db	231	AMACILITGAL	241							
RESULT	12									
Q9V51		PRELIMINARY;	PRT;	386 AA.						
ID	Q9V51									
AC	Q9V51; Q9V50;									
DT	01-MAY-2000 (TremBrel. 13, Created)									
DT	01-MAY-2000 (TremBrel. 13, Last sequence update)									
DT	01-JUN-2001 (TremBrel. 17, Last annotation update)									
DE	CG4729 PROTEIN.									
GN	CG4729.									
OS	Drosophila melanogaster (Fruit fly).									
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Ephyroidae; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha.									
OC	Ephyroidae; Drosophilidae; Drosophila.									
OC	NCBI_Taxid=7227;									
RN	[1]									
RP	SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).									
RC	SPRAIN-BERKELEY.									
RX	MEDLINE-20196006; PubMed-10731132;									
RA	Adams M.D., Celiniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sultion G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D., Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G., Abrell J.F., Agdayan A., An H.-J., Andrews-Plamkoc C., Baldwin D., Bailly R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S., Borokva D., Botchan M.R., Bouck J., Brokstein P., Brotlier P., Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson R., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W., Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K., Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwa C., Jalali M., Kalush F., Kapen G.H., Ke Z., Kennison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Mishaeva N.V., Mobarry C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.N., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M., Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Slier E., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., Spier B.C., Spradling A.C., Stapleton M., Strong R., Sun E., Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.-Y., Weissman D.A., Weinstock G.M., Weissbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.; "The genome sequence of Drosophila melanogaster."; Science 287:2185-2195(2000).									
RT	-1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A SHORT FORM: ARE PRODUCED BY ALTERNATIVE SPLICING.									
CC	EMBL: AB003527; AAF49472.1; -									
DR	EMBL: AEO03527; AAF49472.1; -									
DR	FLYBase: Fgn0036623; CG4729.									
DR	Interpro: IPR002123; Acyltransferase.									
DR	Pfam: PF01553; Acyltransferase; 1.									
FT	Alternative splicing: Hypothetical protein.									
VS	VARSPLIC 1 243 MISSING (IN SHORT ISOFORM).									
SO	SEQUENCE 386 AA; 44397 MW; IE94370E971C35D5 CRC64;									

```

Query Match      14.6%; Score 272.5; DB 5; Length 386;
Best Local Similarity 26.6%; Pred. No. 5.6e-17;
Matches 72; Conservative 63; Mismatches 105; Indels 31; Gaps 9;

QY 35 FYQALDEL-----YCYQ--SMVLFPEFNTGVQILLYGDLP-----KKNENIYLAN 81
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 38 FIKPIDKRLFKRLMYACYSLYSQILFVSDWYAGSKMTYMDKEDFEKHAGHEHLLMN 97
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 82 HQSTVDWIVADILAIRONALGHVRVYLKGLKMLPL--YCGFAGHGCIYKRSKFNEMK 139
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 98 HKYELIDWLNMGWICKLGVLCNCKAYAKKAIKYVPIIGWMLAEF--VFLRNRPDOKT 155
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 140 EMRNKLQ---SYVDAGTDMVLIPEEGTGRYNEBOTKVLASQAFAORGLAVLKHVLP 196
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 156 IITBOLKVVFSYPD---FTWILLNAEGTRFTPAKHE---ASVKFAOEGMIVLAKHLLIP 209
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 197 IKATHVAFDCMKNYLDAYDVTVYVEGKDGQRRSEPTMTTEFLCKECPKIHIDRIDK 256
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 210 TKGFVSLAPIRGLCPVLYIDINLAVRPD-----KTPADMLSLHGKSVEPHLMRRILPL 264
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 257 KDVEPEOEHMRRLHREFEIKDKMLIEYES 287
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 265 EOVPEDEKEAAMLONLEFVEKDKIIDSFLET 295
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 13
0924S1 PRELIMINARY; PRT; 378 AA.
AC 0924S1;
DT 01-DEC-2001 (TREMBlrel. 19, created)
DT 01-DEC-2001 (TREMBlrel. 19, last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, last annotation update)
DE LYSDPHOSPHATIDIC ACID ACYLTRANSFERASE-DELTA.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
OX
RN [1]
RP SEQUENCE FROM N.A.
RA Li W., Suzuki T.,
RT "Rattus norvegicus mRNA for lysophosphatidic acid acyltransferase-
RT delta, complete cds."
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL, AB067572; BAB62290.1;
KW Transferase; Acyltransferase.
SQ SEQUENCE 378 AA; 43794 MW; 389AA01B7327AE2B CRC64;

```

```

Query Match      14.3%; Score 267; DB 5; Length 380;
Best Local Similarity 25.9%; Pred. No. 1.8e-16;
Matches 83; Conservative 69; Mismatches 123; Indels 46; Gaps 9;

QY 25 RLISAFDLARFQALDDRLICYQSMVLFPEFNTGVQILLYGDLPKN-----KENIY 78
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 43 RSLAYLHYFSYCI-----LVCV-----AEWYAGSKRLRYID-PODEKFFGKEHGILL 89
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 14
Q9YV49 PRELIMINARY; PRT; 380 AA.
AC Q9YV49;
DT 01-MAY-2000 (TREMBlrel. 13, created)
DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, last annotation update)
DE CG4753 PROTEIN.
GN CG4753.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
OX
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-BERKELEY;
RC MEDLINE=20196006; PubMed=10731132;
RX Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanosides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt E., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotlier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Dudin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
RA Foeller C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Hounck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Iodagan C.,
RA Jatali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lal Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclio J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL; AE003527; AAF9473.1;
DR FlyBase; FBgn0036622; CG4753.
DR InterPro; IPR002123; Acyltransferase.
DR Pfam; PF01553; Acyltransferase; 1.
SQ SEQUENCE 380 AA; 43482 MW; EA240947566A841 CRC64;

```

[illegible]

```

Db      286 FT-----KDFESDLEVHQINRPKPLIVL-----IMGLVF 318
               | | | | |
QY      341 GTLGLVWFIKA 353
               | | | | |
Db      319 GGFKLQWLSTVA 331

```

Search completed: August 28, 2002, 11:16:55  
Job time: 559 sec

RESULT	15			
99SYC8				
ID	09SYC8	PRELIMINARY;	PRT;	376 AA.
AC	09SYC8;			
DT	01-MAY-2000 (TREMBLrel, 13, Created)			
DT	01-MAY-2000 (TREMBLrel, 13, Last sequence update)			
DT	01-JUN-2001 (TREMBLrel, 17, Last annotation update)			
DE	PURATIVE ACYL-COA:1-ACYLGlycerol-3-PHOSPHATE ACYLTRANSFERASE.			
GN	FLM15.12.			
OS	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eumariota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:			
OC	Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae:			
OC	eucosids II: Brassicales: Brassicaceae: Arabidopsis.			
OX	NCBI_TaxId=3702:			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Federpiet N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,			
RA	Altafi H., Araujo R., Hultzer L., Rowley D., Buehler E., Dunn P.,			
RA	Gonzalez A., Kremetskaia I., Kim C., Lenz C., Li J., Liu S.,			
RA	Luros S., Schwartz J., Shinn P., Tortilini M., Vysotskaia V.S.,			
RA	Walker M., Yu G., Ecker J., Theologis A., Davis R.W.,			
RL	Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AC006085; AAC03638.1; -			
DR	InterPro; IPR002123; ACyltransferase.			
DR	InterPro; IPR003015; HLH_MYC.			
DR	Pfam; PF01553; ACyltransferase; 1.			
DR	PROSITE; PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.			
DR	Transferrase; ACyltransferase.			
QO	SEQUENCE 376 AA; 43435 MW; 0B9AE1A3B09ED4D0 CRC64;			

Query Match	14.18;	Score 263;	DB 10;	Length 376;
Best Local Similarity	23.08;	Pred. No. 4.1e-16;		
Matches	72;	Conservative	76;	Mismatches 117;
			Indels	48;
			Gaps	8

[illegible]





**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:18:44 ; Search time 75.33 Seconds

(without alignments)  
336.717 Million cell updates/sec

Title: US-09-853-526-5

Perfect score: 364  
Sequence: 1 MLSLVLYHTSMRLLPSVY.....YVNWVYGLGLGWTRIKA 364

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 747574 seqs, 111073796 residues

Word size : 0

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

A.Geneseq\_032802:\*  
1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:\*  
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:\*  
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	364	100.0	364	20	AAV36729
2	364	100.0	364	22	AAW50128
3	364	100.0	364	22	AAU00665
4	353	97.0	353	22	AAW85532
5	353	97.0	353	22	AAW85532
6	352	96.7	352	20	AAV36728
7	290	79.7	353	22	AAW39591
8	269	73.9	269	22	AAW09653
9	263	72.3	372	22	AAW41377
10	252	69.2	353	22	AAW93595
11	230	63.2	291	20	AAV36744

12	228	62.6	228	20	AAV36740	Truncated pgl prot
13	199	54.7	261	20	AAV36745	pgl splice variant
14	185	50.8	185	20	AAV36753	pgl splice variant
15	185	50.8	315	20	AAV36751	pgl splice variant
16	184	50.5	300	20	AAV36752	pgl splice variant
17	154	42.3	182	20	AAV36750	pgl splice variant
18	116	31.9	238	20	AAV36743	pgl splice variant
19	85	23.4	96	20	AAV36749	pgl splice variant
20	70	19.2	257	22	ABG01360	Novel human diago
21	70	19.2	980	22	ABG11977	Novel human diago
22	62	17.0	66	20	AAV36748	pgl splice variant
23	62	17.0	68	20	AAV36747	pgl splice variant
24	62	17.0	77	20	AAV36742	pgl splice variant
25	62	17.0	90	20	AAV36746	pgl splice variant
26	53	14.6	97	22	ABG12868	Novel human diago
27	53	14.6	1104	22	ABG12107	Novel human diago
28	32	8.8	354	20	AAV36741	Mouse pgl protein
29	9	2.5	9	20	AAV36734	pgl box 3 motif.
30	9	2.5	157	22	AAW550130	Lysophosphatidic a
31	9	2.5	237	22	AAW55540	H. pylori ORF 12g
32	9	2.5	240	19	AAW98477	H. pylori GRPO 903
33	9	2.5	242	18	AAW35441	H. pylori ORF 069p
34	8	2.2	9	20	AAV36737	Acyl glycerol tran
35	8	2.2	239	22	ABB52710	Escherichia coli p
36	8	2.2	396	21	AAW99484	Yeast acyltransfer
37	8	2.2	557	22	ABB71470	Drosophila melanog
38	8	2.2	1820	22	ABB62370	Drosophila melanog
39	7	1.9	12	16	AAW87705	Cocount Ag3PAT cry
40	7	1.9	12	16	AAW43149	Cocount 1-acylglyc
41	7	1.9	65	16	AAW87720	Cocount LPAT clon
42	7	1.9	65	20	AAW43164	1-acylglycerol-3-p
43	7	1.9	82	16	AAW87718	Cocount LPAT clon
44	7	1.9	82	16	AAW87719	Cocount LPAT clon
45	7	1.9	82	20	AAV43160	1-acylglycerol-3-p

#### ALIGNMENTS

RESULT 1	
AAV36729	standard; Protein: 364 AA.
AC	AAV36729;
DT	27-SEP-1999 (first entry)
DE	Human pgl protein sequence.
XX	
KW	PGI gene; diallelic marker; human; PSA: PGI-related diallelic marker;
KW	cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.
XX	
OS	Homo sapiens.
XX	
PN	MO9932644-A2.
XX	
PD	01-JUL-1999.
XX	
PF	22-DEC-1998; 98WO-1B02133.
XX	
PR	09-SEP-1998; 98US-0099658.
PR	22-DEC-1997; 97US-0996306.
XX	
PA	(GEST) GENSET.
PI	Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;
DR	WPI: 1999-405178/34.
DR	N-PSDB: AA200870.
XX	
PT	Use of a prostate cancer associated gene and diallelic markers
XX	derived from it

PS Claim 7; Page 190-191; 365bp; English.

XX The invention relates to a mammalian PGI gene and protein, and a set of

CC PGI allelic markers. The PGI polynucleotide and allelic markers are

CC used in a hybridisation assay, a sequencing assay, or in an

CC allele-specific amplification assay for determining the identity of a

CC nucleotide at a PGI-related allelic marker. The methods can be used to

CC detect and to assess the risk of developing cancer or prostate cancer.

CC Early-stage diagnosis of prostate cancer relies on prostate specific

CC antigen (PSA) dosage. However, the effectiveness of this is limited due

CC to its inability to discriminate between malignant and non-malignant

CC affections of the organ. A need exists for both a reliable diagnostic

CC procedure which would enable early-stage diagnosis, and for preventative

CC and curative treatments of the disease. The PGI gene can be used for

CC detection of prostate cancer, and the risk of developing it in the

CC future, and can also be used to determine therapies for the disease.

XX

SO Sequence 364 AA:

Query Match 100.0%; Score 364; DB 20; Length 364;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLSTLVHTTSMKRYLPSVYVLGTAPTYVLAWGVWLLSAFLPARYQALDRLCYQOS 60

DB 1 MLSTLVHTTSMRYLPSVYVLGTAPTYVLAWGVWLLSAFLPARYQALDRLCYQOS 60

QY 61 MVLFFPNYNGVOILYGLDLPKKNKENTITVLANHSTVDWIVADTLATRONALGHVRVVK 120

DB 61 MVLFFPNYNGVOILYGLDLPKKNKENTITVLANHSTVDWIVADTLATRONALGHVRVVK 120

QY 121 EGLKMLPLCYCYFAOHGCIYVVKRSKAFNEKEMRNKLOSVDAGTPMYLVIFPESTRNPE 180

DB 121 EGLKMLPLCYCYFAOHGCIYVVKRSKAFNEKEMRNKLOSVDAGTPMYLVIFPESTRNPE 180

QY 181 QTKVLSAQFAAORGLAVLKHVLTPIKATVHAFDCMKVTLDAIVYTVVYEGSKDDGQ 240

DB 181 QTKVLSAQFAAORGLAVLKHVLTPIKATVHAFDCMKVTLDAIVYTVVYEGSKDDGQ 240

QY 241 RRESPTWTEFLCKECPRIHIDRIIDKDVPEDEHMRMLHEFEIKDKMLIEFYSPD 300

DB 241 RRESPTWTEFLCKECPRIHIDRIIDKDVPEDEHMRMLHEFEIKDKMLIEFYSPD 300

QY 301 PERRRRPGRSVNSKLSIKRTPSMLISGLTAGMLMTDAGRKLVTWTVYGTLLGCLMW 360

DB 301 PERRRRPGRSVNSKLSIKRTPSMLISGLTAGMLMTDAGRKLVTWTVYGTLLGCLMW 360

QY 361 TTKA 364

DB 361 TTKA 364

RESULT 2

AAM50128

ID AAM50128 standard; Protein; 364 AA.

XX

AC AAM50128;

XX

DT 21-DEC-2001 (first entry)

XX

DE Human acyltransferase 27417.

XX

XX Acyltransferase 27417; human; cell proliferation; cell migration;

KW cell differentiation; ovary cancer; brain cancer; colon cancer;

KW lung cancer; tumour; metastasis; sarcoma; carcinoma;

KW adenocarcinoma; antitumour; lipid metabolism; diagnosis; therapy.

XX

OS Homo sapiens.

XX

XX

FT Key Location/Qualifiers

FT Peptide 1..49

FT /Label= Signal\_peptide

FT Protein

FT /label= Mature\_protein

FT 50..364

FT Domain

FT /label= N-terminal\_domain

FT 50..320

FT Domain

FT /label= Transmembrane\_domain

FT 321..337

FT Domain

FT /label= C-terminal\_domain

FT 338..364

FT Domain

FT /label= Acyltransferase\_domain

FT 71..363

FT Domain

FT /note= "Asn is N-glycosylated"

FT 68..71

FT Modified-site

FT /note= "O-phosphorylated by CAMP- and

FT 241..244

FT /note= "CGMP-dependent protein kinase"

FT 11..13

FT Modified-site

FT /note= "O-phosphorylated by protein kinase C"

FT 144..146

FT Modified-site

FT /note= "O-phosphorylated by protein kinase C"

FT 205..207

FT Modified-site

FT /note= "O-phosphorylated by protein kinase C"

FT 317..319

FT Modified-site

FT /note= "O-phosphorylated by protein kinase C"

FT 361..363

FT Modified-site

FT /note= "O-phosphorylated by protein kinase C"

FT 95..98

FT Modified-site

FT /note= "O-phosphorylated by casein kinase II"

FT 158..161

FT Modified-site

FT /note= "O-phosphorylated by casein kinase II"

FT 246..249

FT Modified-site

FT /note= "O-phosphorylated by tyrosine kinase"

FT 82..89

FT Modified-site

FT /note= "N-myristoylated"

FT 130..135

FT Modified-site

FT /note= "N-myristoylated"

FT 330..335

FT Modified-site

FT /note= "N-myristoylated"

FT 352..357

FT Modified-site

FT /note= "N-myristoylated"

FT 230..343

FT Modified-site

FT /note= "Amidation"

FT 105..126

FT Peptide

FT /note= "predicted leucine zipper"

PN MO200173051-A2.

XX

XX 04-OCT-2001.

PD

XX

XX 26-MAR-2001; 2001WC-US09633.

PF

XX

PR 24-MAR-2000; 2000US-192092P.

XX

XX (MILL-) MILLENNIUM PHARM INC.

PA

XX

PI Meyers RA, Rudolph-Owen L, Macbeth KJ, Tsai F, Williamson M;

XX

XX WPI; 2001-626267/72.

DR

XX

DR N-PSDB; AAH27094.

XX

XX New polynucleotide encodes a polypeptide for the diagnosis and

PT treatment of cancer and to alter cellular lipid metabolism comprises a

PT polynucleotide encoding human acyltransferases designated 46743 and

PT 27417 -

XX

XX Claim 9; Fig 8A-B; 150bp; English.

PS

XX

XX The present sequence is that of a novel human protein, termed

CC 27417, which shows the structural characteristics of members of the

CC acyltransferase family, including an acyltransferase domain. The

CC invention provides novel acyltransferase 27417 nucleic acids and  
 CC polypeptides, as well as methods for detecting their presence, and  
 CC methods for screening for compounds that modulate their expression  
 CC or activity. Such compounds can be used to treat conditions  
 CC related to aberrant activity or expression of 27417 protein or  
 CC nucleic acid, such as conditions involving deficient cellular  
 CC proliferation, migration and/or differentiation. Also provided are  
 CC methods for inhibiting the proliferation or migration, or inducing  
 CC the killing, of a 27417-expressing cell, e.g. a hyperproliferative  
 CC and/or metastatic cell, by contact with a compound that modulates  
 CC the activity or expression of the 27417 protein or nucleic acid.  
 CC The 27417-expressing cell is found in the healthy or diseased  
 CC heart, blood vessels, kidney, skeletal muscle, brain or liver, or  
 CC especially in a solid tumour, a soft tissue tumour or a metastatic  
 CC lesion, a sarcoma, a carcinoma or an adenocarcinoma, and in  
 CC particular in a hyperproliferative and/or metastatic cell found in  
 CC ovarian, brain, colon or lung cancer. The compounds may also be  
 CC used to modulate lipid metabolism in a 27417-expressing cell.  
 CC Methods of disease diagnosis, e.g. by determining the  
 CC presence of a genetic alteration in a 27417 polypeptide, and for  
 CC evaluating the efficacy of a treatment of a disorder, are also  
 CC provided.

XX Sequence 364 AA;

Query Match 100.0%; Score 364; DB 22; Length 364;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLISLVLHRTYSMRVLLPSVVLGTAFTYVLANGVWRLLSAFLPARYQALDRLXCVYGS 60  
 DB 1 MLISLVLHRTYSMRVLLPSVVLGTAFTYVLANGVWRLLSAFLPARYQALDRLXCVYGS 60  
 QY 61 MVLFFENTGVQILLYGDLPKNKENIIVLANHQSVDVIVADIARONALGHVRYLK 120  
 DB 61 MVLFFENTGVQILLYGDLPKNKENIIVLANHQSVDVIVADIARONALGHVRYLK 120  
 QY 121 EGLKMLPIVGCYFAOAGGIYVKRSKAFNEKEMNKLOSVDAGTPMYLVIFPGSTRNPE 180  
 DB 121 EGLKMLPIVGCYFAOAGGIYVKRSKAFNEKEMNKLOSVDAGTPMYLVIFPGSTRNPE 180  
 QY 181 QTKVLSAQAFAAQRGLAVLKHVLTPIRIKATHVAFDCKMNYLDIAIVTVVYEGKDDGQ 240  
 DB 181 QTKVLSAQAFAAQRGLAVLKHVLTPIRIKATHVAFDCKMNYLDIAIVTVVYEGKDDGQ 240  
 QY 241 RRSPTMTFELCKECPKIHIDRIKDDVPEQEHMRRLHREFIKDKMLIEFYSPD 300  
 DB 241 RRSPTMTFELCKECPKIHIDRIKDDVPEQEHMRRLHREFIKDKMLIEFYSPD 300  
 QY 301 PERRRRPPGKSVNSKISIKKTLPSMLILSGLTAGMLMTDAGRKLVTYNTWYTGLLGLMW 360  
 DB 301 PERRRRPPGKSVNSKISIKKTLPSMLILSGLTAGMLMTDAGRKLVTYNTWYTGLLGLMW 360  
 QY 361 TIRKA 364  
 DB 361 TIRKA 364

RESULT 3  
 AAU00665  
 ID AAU00665 standard; Protein; 364 AA.

XX AAU00665;

XX 07-SEP-2001 (first entry)

DE Human lysophosphatidic acid acyltransferase isoform LPAAT-epsilon.

KW lysophosphatidic acid acyltransferase epsilon; LPAAT-epsilon; LPA; PA;  
 KW lysophosphatidic acid; phosphatidic acid; acylation; cellular activation;  
 KW phospholipid signaling; mitogenesis; inflammation; autoimmune disease;  
 KW oncology; cancer; obesity; gene therapy.

XX OS Homo sapiens.  
 XX PN WO200134782-A1.  
 XX PD 17-MAY-2001.  
 XX PF 02-NOV-2000; 2000WO-US30193.  
 XX PR 09-NOV-1999; 99US-0436919.  
 XX PA (CELL-) CELL THERAPEUTICS INC.  
 XX PI Leung DW;  
 XX DR WPI; 2001-335920/35.  
 XX DR N-PSDB; AAS00649.

PT Novel isolated human isoform of lysophosphatidic acid  
 PT acyltransferase-epsilon useful for diagnostic, therapeutic and  
 PT screening purposes

PS Claim 1; Fig 1; 48pp; English.

XX The sequence represents a human lysophosphatidic acid acyltransferase  
 CC (LPAAT) isoform, LPAAT-epsilon. LPAAT catalyses the acylation of  
 CC lysophosphatidic acid (LPA) to phosphatidic acid (PA). LPA and PA have  
 CC been identified as phospholipid signalling molecules that affect a wide  
 CC range of biological responses. PA is involved in cellular activation and  
 CC mitogenesis. Compounds that block PA generation and hence diminish lipid  
 CC biosynthesis and the signal involved in cell activation are of  
 CC therapeutic interest in the areas of inflammation and oncology (e.g.  
 CC autoimmune diseases and cancer) as well as obesity treatment.  
 CC LPAAT-epsilon and its corresponding DNA can be used in screening assays  
 CC to detect agents that stimulate or inhibit the activity of LPAAT and,  
 CC therefore, PA. The DNA is useful in tests to detect the presence or  
 CC expression of LPAAT-epsilon in relation to certain diseases and  
 CC conditions, and in disease prevention and treatment. The sequences of the  
 CC invention are also useful for diagnosis of diseases and conditions in  
 CC which the expression of LPAAT enzyme is abnormal.

SQ Sequence 364 AA;

Query Match 100.0%; Score 364; DB 22; Length 364;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLISLVLHRTYSMRVLLPSVVLGTAFTYVLANGVWRLLSAFLPARYQALDRLXCVYGS 60  
 DB 1 MLISLVLHRTYSMRVLLPSVVLGTAFTYVLANGVWRLLSAFLPARYQALDRLXCVYGS 60  
 QY 61 MVLFFENTGVQILLYGDLPKNKENIIVLANHQSVDVIVADIARONALGHVRYLK 120  
 DB 61 MVLFFENTGVQILLYGDLPKNKENIIVLANHQSVDVIVADIARONALGHVRYLK 120  
 QY 121 EGLKMLPIVGCYFAOAGGIYVKRSKAFNEKEMNKLOSVDAGTPMYLVIFPGSTRNPE 180  
 DB 121 EGLKMLPIVGCYFAOAGGIYVKRSKAFNEKEMNKLOSVDAGTPMYLVIFPGSTRNPE 180  
 QY 181 QTKVLSAQAFAAQRGLAVLKHVLTPIRIKATHVAFDCKMNYLDIAIVTVVYEGKDDGQ 240  
 DB 181 QTKVLSAQAFAAQRGLAVLKHVLTPIRIKATHVAFDCKMNYLDIAIVTVVYEGKDDGQ 240  
 QY 241 RRSPTMTFELCKECPKIHIDRIKDDVPEQEHMRRLHREFIKDKMLIEFYSPD 300  
 DB 241 RRSPTMTFELCKECPKIHIDRIKDDVPEQEHMRRLHREFIKDKMLIEFYSPD 300  
 QY 301 PERRRRPPGKSVNSKISIKKTLPSMLILSGLTAGMLMTDAGRKLVTYNTWYTGLLGLMW 360  
 DB 301 PERRRRPPGKSVNSKISIKKTLPSMLILSGLTAGMLMTDAGRKLVTYNTWYTGLLGLMW 360  
 QY 361 TIRKA 364

||||  
Db 361 tika 364  
RESULT 4  
ID AAB85532 standard; protein; 353 AA.  
XX AAB85532:  
XX 25-SEP-2001 (first entry)  
XX Human secreted protein (clone id HLTH084).  
XX Secreted protein; immunosuppressive; antiarthritic; antirheumatic;  
KW antiproliferative; cytosolic; cardiant; vasotropic; cerebroprotective;  
KW neotropic; neuroprotective; antibacterial; virucide; fungicide; human;  
KM opticalmalogical; gene therapy.  
XX Homo sapiens.  
XX MO200155430-A1.  
XX 02-AUG-2001.  
XX 17-JAN-2001: 2001MO-US01431.  
XX 31-JAN-2000: 2000US-0179065.  
PR 04-FEB-2000: 2000US-0180628.  
PR 12-SEP-2000: 2000US-0231968.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR, Olsen HS;  
PI Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH, Fiscella M;  
PI NI J, Ruben SM, Barash SC;  
XX WPI: 2001-476220/51.  
DR N-PSDB: AAAH46942.  
XX  
XX 17 Isolated nucleic acid molecules encoding human secreted proteins,  
PT used to preventing, treating or ameliorating a medical condition -  
XX  
XX Claim 11: Page 450-451; 482pp; English.  
XX  
XX The invention provides novel human secreted proteins and polynucleotides  
CC encoding them. The secreted proteins can be expressed by standard  
CC recombinant methodology. The secreted proteins and polynucleotides are  
CC used to prevent, treat or ameliorate a medical condition in e.g. humans,  
CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They can  
CC also be used in diagnosing a pathological condition. The antibodies to  
CC the proteins can also be used in alleviating symptoms associated with the  
CC disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme  
CC linked immunosorbent assays (ELISA). Disorders which are diagnosed or  
CC treated include autoimmune diseases e.g. rheumatoid arthritis,  
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,  
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
CC e.g. cerebral ischemia, anglogenesis, nervous system disorders e.g.  
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and  
CC ocular disorders e.g. corneal infection. The polypeptides can also be  
CC used to aid wound healing and epithelial cell proliferation, to prevent  
CC skin aging due to sunburn, to maintain organs before transplantation, for  
CC supporting cell culture of primary tissues, to regenerate tissues and in  
CC chemotaxis. The polypeptides can also be used as a food additive or  
CC preservative to increase or decrease storage capabilities. The present  
CC sequence represents a human secreted protein.  
XX  
XX Sequence 353 AA:  
Query Match 97.0%; Score 353; DB 22; Length 353;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 MRYLPSVVLGTAPTYVLAMGWRLISATLPARFIOALDDRLKCYQSWLFFPENYTG 71  
Db 1 mrylpsvvlgtaptyvllamgwrlisatlparfioalddrlycqvyswnvllffenytg 60  
QY 72 VOILLGDDLPKNENITYLANNHSTVDWIVADILATRONALGHVRVYLKGLKWLPLYYG 131  
Db 61 vqlllygdldrpknenillyannhstvydwivadilaltronalghvrylvlkeglkwlpllygc 120  
QY 132 YFAOHGIIYVKSRAKFNKEMRNKLOSVDAGTPMYLVLPPEGTRYNPEQTKVLSAQAF 191  
Db 121 yfaohggiyvkrsakfnkemrnklosyvdagtpmylvlppegtrynpeqtkvlsasqaf 180  
QY 192 AAQRGALAVLKHVLTTPRKATNVAFCCKNTLDAIIVDTVYVECKDDGGQRRSPYMERL 251  
Db 181 aadrgalavlkhvlttprkatnva fccckntldaiivdtyvveckddggqrrspymertel 240  
QY 252 CKECPRKHIIHIDRDKDVEDEQEHMRWRWHEFEELKDKDLIFEYSPDERKRRPGKS 311  
Db 241 ckecpkhiihidrdkdveedeqehmrwrwhefeelkdkmli feyspderkrrtpgks 300  
QY 312 VNSKLSIKKTLPSMLILSGLTAGMLTDAGRKLYVNTWYIGTLGLCWYTIKA 364  
Db 301 vnskslkkclpsmllsgltagmltdagrkllyvntwlygtlglcwlwtika 353  
RESULT 5  
ID AAB85552 standard; protein; 353 AA.  
XX AAB85552:  
XX 25-SEP-2001 (first entry)  
XX Human secreted protein (clone id HSLIA81).  
XX  
XX Secreted protein; immunosuppressive; antiarthritic; antirheumatic;  
KW antiproliferative; cytosolic; cardiant; vasotropic; cerebroprotective;  
KW neotropic; neuroprotective; antibacterial; virucide; fungicide; human;  
KM opticalmalogical; gene therapy.  
XX  
XX Homo sapiens.  
XX MO200155430-A1.  
XX 02-AUG-2001.  
XX 17-JAN-2001: 2001MO-US01431.  
XX 31-JAN-2000: 2000US-0179065.  
PR 04-FEB-2000: 2000US-0180628.  
PR 12-SEP-2000: 2000US-0231968.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR, Olsen HS;  
PI Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH, Fiscella M;  
PI NI J, Ruben SM, Barash SC;  
XX WPI: 2001-476220/51.  
DR N-PSDB: AAAH46962.  
XX  
XX 17 Isolated nucleic acid molecules encoding human secreted proteins,  
PT used to preventing, treating or ameliorating a medical condition -  
XX  
XX Claim 11: Page 465-466; 482pp; English.  
XX  
XX The invention provides novel human secreted proteins and polynucleotides  
CC encoding them. The secreted proteins can be expressed by standard  
CC recombinant methodology. The secreted proteins and polynucleotides are  
CC used to prevent, treat or ameliorate a medical condition in e.g. humans,  
CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They can  
CC also be used in diagnosing a pathological condition. The antibodies to

CC the proteins can also be used in alleviating symptoms associated with the  
 CC disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme  
 CC linked immunosorbent assays (ELISA). Disorders which are diagnosed or  
 CC treated include autoimmune diseases e.g. rheumatoid arthritis,  
 CC hyperproliferative disorders e.g. neoplasms of the breast or liver,  
 CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
 CC e.g. cerebral ischemia, angiodenesis, nervous system disorders e.g.  
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and  
 CC ocular disorders e.g. corneal infection. The polypeptides can also be  
 CC used to aid wound healing and epithelial cell proliferation, to prevent  
 CC skin aging due to sunburn, to maintain organs before transplantation, for  
 CC supporting cell culture of primary tissues, to regenerate tissues and in  
 CC chemotaxis. The polypeptides can also be used as a food additive or  
 CC preservative to increase or decrease storage capabilities. The present  
 CC sequence represents a human secreted protein.

XX Sequence 353 AA;

Query Match 97.0%; Score 353; DB 22; Length 353;  
 Best local Similarity 100.0%; Pred. No. 0;  
 Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 MRLTLDPSVLLGTAFTYVLAAGVWRLISAFLEPRFYQALDDRLCYQSWLFEFFENYTG 71  
 DB 1 MYLLPSVLLGTAPYVLAAGVWRLISAFLEPRFYQALDDRLCYQSWLFEFFENYTG 60

QY 72 VOILLVGDLPKKNKENTIIYLANHSTVDWIVADILAIROMLGHVRYLKEGLKWLPLYGC 131  
 DB 61 VQILLVGDLPKKNKENTIIYLANHSTVDWIVADILAIROMLGHVRYLKEGLKWLPLYGC 120

QY 132 YFAOHGIIYKRSRAKFEKEMRNKLOSVDAGTPMYLVPEECTRYNPEOTKVLASQAF 191  
 DB 121 YFQHGIIYKRSRAKFEKEMRNKLOSVDAGTPMYLVPEECTRYNPEOTKVLASQAF 180

QY 192 AAGRGALVLAHVLTPRIKATHVAFDCMKNYLAIYDVTVVYEGKDDGGORRESPTMTFEL 251  
 DB 181 AAGRGALVLAHVLTPRIKATHVAFDCMKNYLAIYDVTVVYEGKDDGGORRESPTMTFEL 240

QY 252 CKCEPKHIIHIDIKDDVEEBOEHMRRLHERFEIKDKMLIEFYSPDERRRKRPFGKS 311  
 DB 241 CKCEPKHIIHIDIKDDVEEBOEHMRRLHERFEIKDKMLIEFYSPDERRRKRPFGKS 300

QY 312 VNSKLSIKKTLPMSLLISGLTAGMLMTDAGRKLIVNTWIVYGLLGLCLWTVIK 364  
 DB 301 VNSKLSIKKTLPMSLLISGLTAGMLMTDAGRKLIVNTWIVYGLLGLCLWTVIK 353

# RESULT 6

AA36728  
 ID AAY36728 standard; Protein; 352 AA.

AC AAY36728;

DT 27-SEP-1999 (first entry)

DE Human PGI protein sequence.

XX PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;  
 KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.  
 XX Homo sapiens.

XX WO932644-A2.

XX 01-JUL-1999.

XX 22-DEC-1998; 98WO-1B02133.

XX 09-SEP-1998; 98US-0099658.  
 PR 22-DEC-1997; 97US-0996506.

XX (GEST ) GENSET.

XX Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;  
 PI MPI: 1999-405178/34.  
 DR N-PSDB: AA200872.

XX Use of a prostate cancer associated gene and biallelic markers  
 PT derived from it  
 PS Claim 7; Page 190; 385pp; English.

CC The invention relates to a mammalian PGI gene and protein, and a set of  
 CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
 CC used in a hybridisation assay, a sequencing assay, or in an  
 CC allele-specific amplification assay for determining the identity of a  
 CC nucleotide at a PGI-related biallelic marker. The methods can be used to  
 CC detect and to assess the risk of developing cancer or prostate cancer.  
 CC Early-stage diagnosis of prostate cancer relies on prostate specific  
 CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
 CC to its inability to discriminate between malignant and non-malignant  
 CC affections of the organ. A need exists for both a reliable diagnostic  
 CC procedure which would enable early-stage diagnosis, and for preventive  
 CC and curative treatments of the disease. The PGI gene can be used for  
 CC detection of prostate cancer, and the risk of developing it in the  
 CC future, and can also be used to determine therapies for the disease.

SO Sequence 352 AA;

Query Match 96.7%; Score 352; DB 20; Length 352;  
 Best local Similarity 100.0%; Pred. No. 0;  
 Matches 352; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 MRLTLDPSVLLGTAFTYVLAAGVWRLISAFLEPRFYQALDDRLCYQSWLFEFFENYTG 71  
 DB 1 MYLLPSVLLGTAPYVLAAGVWRLISAFLEPRFYQALDDRLCYQSWLFEFFENYTG 60

QY 72 VOILLVGDLPKKNKENTIIYLANHSTVDWIVADILAIROMLGHVRYLKEGLKWLPLYGC 131  
 DB 61 VQILLVGDLPKKNKENTIIYLANHSTVDWIVADILAIROMLGHVRYLKEGLKWLPLYGC 120

QY 132 YFAOHGIIYKRSRAKFEKEMRNKLOSVDAGTPMYLVPEECTRYNPEOTKVLASQAF 191  
 DB 121 YFQHGIIYKRSRAKFEKEMRNKLOSVDAGTPMYLVPEECTRYNPEOTKVLASQAF 180

QY 192 AAGRGALVLAHVLTPRIKATHVAFDCMKNYLAIYDVTVVYEGKDDGGORRESPTMTFEL 251  
 DB 181 AAGRGALVLAHVLTPRIKATHVAFDCMKNYLAIYDVTVVYEGKDDGGORRESPTMTFEL 240

QY 252 CKCEPKHIIHIDIKDDVEEBOEHMRRLHERFEIKDKMLIEFYSPDERRRKRPFGKS 311  
 DB 241 CKCEPKHIIHIDIKDDVEEBOEHMRRLHERFEIKDKMLIEFYSPDERRRKRPFGKS 300

QY 312 VNSKLSIKKTLPMSLLISGLTAGMLMTDAGRKLIVNTWIVYGLLGLCLWTVIK 363  
 DB 301 VNSKLSIKKTLPMSLLISGLTAGMLMTDAGRKLIVNTWIVYGLLGLCLWTVIK 352

# RESULT 7

AA39591  
 ID AAM39591 standard; Protein; 353 AA.

AC AAM39591;

DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 2736.

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;





PR	17-NOV-2000;	2000US-0249216.
PR	17-NOV-2000;	2000US-0249217.
PR	17-NOV-2000;	2000US-0249218.
PR	17-NOV-2000;	2000US-0249219.
PR	17-NOV-2000;	2000US-0249220.
PR	17-NOV-2000;	2000US-0249221.
PR	17-NOV-2000;	2000US-0249222.
PR	17-NOV-2000;	2000US-0249223.
PR	17-NOV-2000;	2000US-0249224.
PR	17-NOV-2000;	2000US-0249225.
PR	17-NOV-2000;	2000US-0249226.
PR	17-NOV-2000;	2000US-0249227.
PR	17-NOV-2000;	2000US-0249228.
PR	17-NOV-2000;	2000US-0249229.
PR	17-NOV-2000;	2000US-0249230.
PR	01-DEC-2000;	2000US-0250160.
PR	01-DEC-2000;	2000US-0250391.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251988.
PR	05-DEC-2000;	2000US-0256719.
PR	06-DEC-2000;	2000US-0251479.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251868.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251989.
PR	08-DEC-2000;	2000US-0251990.
PR	11-DEC-2000;	2000US-0254097.
PR	05-JAN-2001;	2001US-0259678.
XX		
PA	(HUMA-) HUMAN GENOME SCI INC.	
PI	Rosen CA, Barash SC, Ruben SM;	
XX		
DR	WPI; 2001-457716/49.	
DR	N-PSDB; AAD16620.	
XX		
PT	Isolated lipid metabolism polypeptide for screening to identify	
PT	antagonists and agonists that may enhance or block activities mediated	
PT	by lipid metabolism proteins and also for testing and detection e.g.	
PT	diagnosis -	
XX		
PS	Claim 11; SEQ ID NO 47; 453pp; English.	
XX		
CC	AAD16606-AAD16627 represent cDNAs corresponding to human lipid metabolism	
CC	protein genes, and AAEO9639-AAEO9660 represent the proteins they encode.	
CC	AAD16628-AAD16647 represent human lipid metabolism genomic DNAs. The	
CC	genes and their corresponding proteins are useful for preventing,	
CC	treatment or ameliorating medical conditions such as immune disorders,	
CC	inflammatory disorders (e.g., conjunctivitis, glossitis), autoimmune	
CC	diseases (e.g., rheumatoid arthritis, systemic lupus erythematosus),	
CC	respiratory disorders (asthma, allergy), CNS disorders (e.g., Alzheimer's	
CC	disease, Parkinson's disease), haematopoietic disorders, cardiovascular	
CC	disorders (e.g., atherosclerosis, coronary disease), hyperlipidaemias,	
CC	hyperlipidaemia, lipidoses, familial hypercholesterolaemia, Gaucher's	
CC	disease, Tay-Sachs's disease, diabetes mellitus, cancer and other enzyme	
CC	abnormalities. They are used for the prevention of mental retardation	
CC	and the enlargement of spleen and liver. They are used in gene therapy	
CC	and antisense therapy. The present sequence represents a human lipid	
CC	metabolism protein of the invention.	
CC	Note: The sequence data for this patent did not form part of the printed	
CC	specification, but was obtained in electronic format directly from WIPO	
CC	at ftp.wipo.int/pub/published_pct_sequences.	
XX		
SQ	Sequence 269 AA;	
	Query Match 73.9%; Score 269; DB 22; Length 269;	
	Best Local Similarity 100.0%; Pred. No. 1,6e+264;	
	Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	TVDVTADIIILIRONALGHRYVLKEGKWLPLVGCGYFAOHGGIYVKRSAPNEKEMRNK 155	
	.....	
Dd	1 tvdtvtadililrinalghrvyvlkegkwlplvgcgyfaohggilyvkrsafnekenrnk 60	
	.....	
OY	LQSYVDAGTPEMYLYIFPEGTRYNEDQTKVLSASQAFAAQRGLAYLKHVLPRIKATHVAF 215	
	.....	
Dd	lqsyvdaqtpeymlyifpegtrynedqtkvlsasqaafaqrqlaylkhlvprikathvaf 120	

QY 216 DCKKNYDAIYDTVYVEGDGQRESEPTMEFECKECPKIHIDRIDKDVPEEQE 275  
CC |  
CC |  
Db 121 dcmynydaIydvtyvegkddgqgrrspcmteflckepkIhIdridkdvpege 180  
QY 276 HMRRLHEREFEIKDKMLIEFYESPDPERRRKRPFGKSVNSKLSIKKTLPSMLISGLTAGM 335  
Db 181 hmrwrlherfelkdkmliefyespdpertrrfpgksvnsklsikklpsmlisgltagm 240  
QY 336 LMTDAGRRLVNTWYVYGTLLGCLMWITKA 364  
Db 241 lmtdagrklyvntwiygtlllgclmwitka 269  
RESULT 9  
ID AAM41377 standard; Protein: 372 AA.  
XX AAM41377;  
AC AAM41377;  
XX 22-OCT-2001 (first entry)  
DT Human polypeptide SEQ ID NO 6308.  
XX  
DE Human: nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
XX peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokine; thrombolytic; drug screening; arthritis; inflammation;  
KW Leukemia.  
XX  
OS Homo sapiens.  
XX  
PN WO200153312-A1.  
PD 26-JUL-2001.  
XX  
PF 26-DEC-2000; 2000MO-US34263.  
XX  
PR 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-0552317.  
PR 09-JUL-2000; 2000US-0598042.  
PR 19-JUL-2000; 2000US-0620312.  
PR 03-AUG-2000; 2000US-0653450.  
PR 14-SEP-2000; 2000US-0662191.  
PR 19-OCT-2000; 2000US-0693036.  
PR 29-NOV-2000; 2000US-0727344.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX  
DR WPI: 2001-442253/47.  
DR N-PSDB: AA160553.  
XX  
XX Novel nucleic acids and polypeptides, useful for treating disorders  
PT such as central nervous system injuries -  
XX  
PS Example 2; SEQ ID NO 6308; 10078bp; English.  
XX  
CC The invention relates to human nucleic acids (AA157798-AA161369) and  
CC the encoded polypeptides (AAM38642-AA442213) with nootropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,

CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.  
XX  
SQ Sequence 372 AA:  
QY 1 MLSLVHTYSMKRLPSPVVLGTAFTYVLAWGVRLSAFLPAREYQALDRLYCVYGS 60  
Db 9 mlslvhtysmryllpsvvlgtapfyvawgvrlsaflparfygaldrlycvygs 68  
QY 61 MVLFFENYGVQIILYGDLPKKNKENTIYLANOSYVDWIVADILAIROMALGHVRVYK 120  
Db 69 mvlffenygvqilliygdlpkknkenillyanqstvdwlvadilalrqnalghvrvyk 128  
QY 121 EGLKWLPLVGCYFAOHGCIYVVKRSKAKENKEMRNKLSQSYVDAGTPLYVIFPGSTRYNP 180  
Db 129 eglkwlplygcyfaqhgciyvkrksakfnekemrnlqsyvdagtpmlyvifpgstrynpe 188  
QY 181 QTRVLSASQAPAAQRCGLAVLKHYLTPRIKATHYAFDCMKNYLDAIYDVYVYEGKDDGQ 240  
Db 189 qtrvlsasqaafaqrglavlkhytlprikatvhafdcmknyldaiydvtyvegkddgq 248  
QY 241 RRESPWTEFLCKECPKIHIDRIDKDVPEEOEHMRRLHBRFEKDKMLIEFESPD 300  
Db 249 rrespwteflckepkIhIdridkdvpeeehmrrlhbrfelkdkmliefespd 308  
QY 301 PERRKRPPGKSVNSKLSIKKTLPSMLISGLTAGMLMTDAGRRLVYNTWYVYGTLLGCLMW 360  
Db 309 perkrtrpgksvnsklsikkltpsmllsgltagmlmtdagrklyvntwiygtllgclmw 368  
QY 361 TIRKA 364  
Db 369 tlrka 372  
RESULT 10  
ID AAB93595 standard; Protein: 353 AA.  
XX AAB93595;  
AC AAB93595;  
XX  
DT 26-JUN-2001 (first entry)  
XX  
DE Human protein sequence SEQ ID NO:13028.  
XX  
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.  
XX  
OS Homo sapiens.  
XX  
PN EP1074617-A2.  
PD 07-FEB-2001.  
XX  
PF 28-JUL-2000; 2000EP-0116126.  
XX  
PR 29-JUL-1999; 99JP-0248036.  
PR 27-AUG-1999; 99JP-0300253.  
PR 11-JAN-2000; 2000JP-0118776.  
PR 02-MAY-2000; 2000JP-0183767.  
PR 09-JUN-2000; 2000JP-0241899.  
XX  
PA (HELI-) HELIX RES INST.  
XX  
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
XX  
DR WPI: 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602  
PT full-length cDNAs defined in the specification, and for the detection  
PT and/or diagnosis of the abnormality of the proteins encoded by the  
PT full-length cDNAs -

PS Claim 8: SEQ ID 13028; 2537bp + CD ROM; English.

CC The present invention describes primer sets for synthesizing 5602  
CC full-length cDNAs defined in the specification. Where a primer set  
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
CC to the complementary strand of a polynucleotide which comprises one of  
CC the 5602 nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in  
CC the specification. The primer sets can be used in antisense therapy and  
CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to  
CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632  
CC represent oligonucleotides, all of which are used in the exemplification  
CC of the present invention.

XX SQ Sequence 353 AA:

Query Match 69.2%; Score 252; DB 22; Length 353;  
Best Local Similarity 99.7%; Pred. No. 3.7e-247;  
Matches 352; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 12 MRLLPSVLLGAPTYVAVGWRLLSAFLPARFYQALDDRLCYVQSVLFFEFNYG 71  
DB 1 MYLLPSVLLGAPTYVAVGWRLLSAFLPARFYQALDDRLCYVQSVLFFEFNYG 60  
OY 72 VOILLYGDLPKKNENIYLANHSTVDWIVADILAIKRONALGHVRYLKEGLMPLLYGC 131  
DB 61 VQILLYGDLPKKNENIYLANHSTVDWIVADILAIKRONALGHVRYLKEGLMPLLYGC 120  
OY 132 YPAHGGIYVKSRAKFNEMRNKLOSVDAGTPTLVTFPESTRNPEOTKVLASQAF 191  
DB 121 YPAHGGIYVKSRAKFNEMRNKLOSVDAGTPTLVTFPESTRNPEOTKVLASQAF 180  
OY 192 AAGRLAVLKHVLPRIKATHVAFDCKMKNYLDIYVTVVYEGSKDDGGORRESPTMTEFL 251  
DB 181 AAGRLAVLKHVLPRIKATHVAFDCKMKNYLDIYVTVVYEGSKDDGGORRESPTMTEFL 240  
OY 252 CKECPRIHIDRIDKDDVEEOEHMRNRLHERFEIKDKMLIEFYSPDERRRKRPFGKS 311  
DB 241 CKECPRIHIDRIDKDDVEEOEHMRNRLHERFEIKDKMLIEFYSPDERRRKRPFGKS 300  
OY 312 VNSKLSIKTLPMSMLILSGITAGMLMTDAGRKLYVNTWYITGLGLCWYTIKA 364  
DB 301 VNSKLSIKTLPMSMLILSGITAGMLMTDAGRKLYVNTWYITGLGLCWYTIKA 353

RESULT 11

ID AAY36744 standard; Protein; 291 AA.

XX AAY36744;

XX 27-SEP-1999 (first entry)

XX PGI splice variant allele protein sequence.

KW PGI gene; diallelic marker; human; PSA; PGI-related diallelic marker;  
KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.  
XX Homo sapiens.

PN WO932644-A2.

XX 01-JUL-1999.

XX 22-DEC-1998; 98WO-1B02133.

XX 09-SEP-1998; 98US-0099658.

XX 22-DEC-1997; 97US-0996306.

XX (GEST ) GENSET.

XX Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;  
PI WPI: 1999-405178/34.

XX N-PSDB; AA00970.

XX Use of a prostate cancer associated gene and diallelic markers  
PT derived from it

PS Claim 7: Page 250-251; 385pp; English.

CC The invention relates to a mammalian PGI gene and protein, and a set of  
CC PGI diallelic markers. The PGI polynucleotide and diallelic markers are  
CC used in a hybridisation assay, a sequencing assay, or in an  
CC allele-specific amplification assay for determining the identity of a  
CC nucleotide at a PGI-related diallelic marker. The methods can be used to  
CC detect and to assess the risk of developing cancer or prostate cancer.  
CC Early-stage diagnosis of prostate cancer relies on prostate specific  
CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
CC to its inability to discriminate between malignant and non-malignant  
CC affections of the organ. A need exists for both a reliable diagnostic  
CC procedure which would enable early-stage diagnosis, and for preventative  
CC and curative treatments of the disease. The PGI gene can be used for  
CC detection of prostate cancer, and the risk of developing it in the  
CC future, and can also be used to determine therapies for the disease.

XX SQ Sequence 291 AA:

Query Match 63.2%; Score 230; DB 20; Length 291;  
Best Local Similarity 100.0%; Pred. No. 6.7e-225;  
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 135 QHGGIYVKSRAKFNEMRNKLOSVDAGTPTLVTFPESTRNPEOTKVLASQAFPAQ 194  
DB 62 QHGGIYVKSRAKFNEMRNKLOSVDAGTPTLVTFPESTRNPEOTKVLASQAFPAQ 121  
OY 195 RGLAVLKHVLPRIKATHVAFDCKMKNYLDIYVTVVYEGSKDDGGORRESPTMTEFL 254  
DB 122 RGLAVLKHVLPRIKATHVAFDCKMKNYLDIYVTVVYEGSKDDGGORRESPTMTEFL 181  
OY 255 CKECPRIHIDRIDKDDVEEOEHMRNRLHERFEIKDKMLIEFYSPDERRRKRPFGKS 314  
DB 182 CKECPRIHIDRIDKDDVEEOEHMRNRLHERFEIKDKMLIEFYSPDERRRKRPFGKS 241  
OY 315 KLSIKTLPMSMLILSGITAGMLMTDAGRKLYVNTWYITGLGLCWYTIKA 364  
DB 242 KLSIKTLPMSMLILSGITAGMLMTDAGRKLYVNTWYITGLGLCWYTIKA 291

RESULT 12

ID AAY36740 standard; Protein; 228 AA.

XX AAY36740;

XX 27-SEP-1999 (first entry)

XX

DE Truncated PGI protein sequence.  
XX  
KW PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;  
XX cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.  
XX  
OS Homo sapiens.  
XX  
PN MO9932644-A2.  
XX  
PD 01-JUL-1999.  
XX  
PE 22-DEC-1998; 98WO-IB02133.  
XX  
PR 09-SEP-1998; 98US-0099658.  
XX 22-DEC-1997; 97US-0996306.  
XX  
PA (GEST ) GENSET.  
XX  
PI Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;  
XX  
DR WPI: 1999-405178/34.  
DR N-PSDB: AA200926.  
XX  
PT Use of a prostate cancer associated gene and biallelic markers  
XX derived from it  
XX  
PS Claim 7: Page 208; 385pp; English.  
XX  
XX The invention relates to a mammalian PGI gene and protein, and a set of  
CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
CC used in a hybridisation assay, a sequencing assay, or in an  
CC allele-specific amplification assay for determining the identity of a  
CC nucleotide at a PGI-related biallelic marker. The methods can be used to  
CC detect and to assess the risk of developing cancer or prostate cancer.  
CC Early-stage diagnosis of prostate cancer relies on prostate specific  
CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
CC to its inability to discriminate between malignant and non-malignant  
CC affections of the organ. A need exists for both a reliable diagnostic  
CC procedure which would enable early-stage diagnosis, and for preventative  
CC and curative treatments of the disease. The PGI gene can be used for  
CC detection of prostate cancer, and the risk of developing it in the  
CC future, and can also be used to determine therapies for the disease.  
XX  
XX Sequence 228 AA:  
SQ

Query Match 62.6%; Score 228; DB 20; Length 228;  
Best Local Similarity 100.0%; Pred. No. 5.7e-223;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 MYLLPSVLLGTAPYVLAMGWRLSAFLPARFYQALDDRLXCVYQSMVLEFFENYTG 71  
DB 1 myllpsvllgtapylvlamgwrlsafflparfyqalddrlxcvysmvleffenytg 60  
QY 72 VQILLKXDLRKKNENTITLNNHSTVDVVIADILATRONALGNRYVLEKGLKPLRYGC 131  
DB 61 vqillkxdlrkknentitlannhstvdvviadilatronalgnryvlekgllkwpilrygc 120  
QY 132 YFQNHGIIYVRSKAKNEKEMRNKLOSIVDAGTBMVIVIEPEGGRYRPEOTKVLASQAF 191  
DB 121 yfqnghiiyvrsaknekeemrnlqsyvdaagtrpmvliwiepgrtyrpeotkvlasaaf 180  
QY 192 AAGRGALVLAHVLPRIKATNVAFDCKKNYLDIYDVTYVVEGKDDG 239  
DB 181 aagrgalviahvlprlkatnvaefdccknyldiydvtvyvegkddgg 228

RESULT 13  
AA36745  
ID AAY36745 standard; Protein; 261 AA.  
XX  
NC AAY36745;  
XX

DT 27-SEP-1999 (first entry)  
XX  
DE PGI splice variant allele protein sequence.  
XX  
KW PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;  
XX cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.  
XX  
OS Homo sapiens.  
XX  
PN MO9932644-A2.  
XX  
PD 01-JUL-1999.  
XX  
PE 22-DEC-1998; 98WO-IB02133.  
XX  
PR 09-SEP-1998; 98US-0099658.  
XX 22-DEC-1997; 97US-0996306.  
XX  
PA (GEST ) GENSET.  
XX  
PI Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;  
XX  
DR WPI: 1999-405178/34.  
DR N-PSDB: AA200971.  
XX  
PT Use of a prostate cancer associated gene and biallelic markers  
XX derived from it  
XX  
PS Claim 7: Page 251; 385pp; English.  
XX  
XX The invention relates to a mammalian PGI gene and protein, and a set of  
CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
CC used in a hybridisation assay, a sequencing assay, or in an  
CC allele-specific amplification assay for determining the identity of a  
CC nucleotide at a PGI-related biallelic marker. The methods can be used to  
CC detect and to assess the risk of developing cancer or prostate cancer.  
CC Early-stage diagnosis of prostate cancer relies on prostate specific  
CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
CC to its inability to discriminate between malignant and non-malignant  
CC affections of the organ. A need exists for both a reliable diagnostic  
CC procedure which would enable early-stage diagnosis, and for preventative  
CC and curative treatments of the disease. The PGI gene can be used for  
CC detection of prostate cancer, and the risk of developing it in the  
CC future, and can also be used to determine therapies for the disease.  
XX  
XX Sequence 261 AA:  
SQ

Query Match 54.7%; Score 199; DB 20; Length 261;  
Best Local Similarity 100.0%; Pred. No. 1.8e-193;  
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 MYLVIEPEGGRYRPEOTKVLASQAFQALVLAHVLPRIKATNVAFDCKKNYLDI 225  
DB 63 mylviepeggrtyrpeotkvlasaafqalviahvlprlkatnvaefdccknyldai 122  
QY 226 YDVTYVVEGKDDGQRRESPTMEFLCKECPKIHIDRIDKQDVEDEGSHMRWLHERF 285  
DB 123 ydvtvyvegkddgqrrresptmeflckecpkihidridkdvpeegshmrwlherf 182  
QY 286 EIKDKMLIERESPDERRRRFPKGSVNSKLSIKTLPMLISGLTAGMLMDAGRKLY 345  
DB 183 eikdkmliefespderrrrfpkgsvnsklsiktlpmlisgltagmlmdagrkly 242  
QY 346 VNTWIGTILGCLMWTIKA 364  
DB 243 vntwigtcilgclmwitika 261

RESULT 14  
AA36753  
ID AAY36753 standard; Protein; 185 AA.  
XX

AC AAY36753;  
XX  
XX 27-SEP-1999 (first entry)  
XX  
DE PGI splice variant allele protein sequence.  
XX  
XX PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;  
KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.  
XX  
OS Homo sapiens.  
XX  
PN WO932644-A2.  
XX  
PD 01-JUL-1999.  
XX  
PF 22-DEC-1998; 98WO-IB02133.  
XX  
XX 09-SEP-1998; 98US-0099658.  
PR 22-DEC-1997; 97US-0996306.  
XX  
XX (GEST ) GENSET.  
XX  
PI Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;  
XX  
XX WPI: 1999-405178/34.  
DR N-PSDB; AA200979.  
XX  
XX Use of a prostate cancer associated gene and biallelic markers  
PT derived from it  
XX  
XX Claim 7; Page 255; 385pp; English.  
XX  
XX The invention relates to a mammalian PGI gene and protein, and a set of  
CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
CC used in a hybridisation assay, a sequencing assay, or in an  
CC allele-specific amplification assay for determining the identity of a  
CC nucleotide at a PGI-related biallelic marker. The methods can be used to  
CC detect and to assess the risk of developing cancer or prostate cancer.  
CC Early-stage diagnosis of prostate cancer relies on prostate specific  
CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
CC to its inability to discriminate between malignant and non-malignant  
CC affections of the organ. A need exists for both a reliable diagnostic  
CC procedure which would enable early-stage diagnosis, and for preventative  
CC and curative treatments of the disease. The PGI gene can be used for  
CC detection of prostate cancer, and the risk of developing it in the  
CC future, and can also be used to determine therapies for the disease.  
XX  
XX Sequence 185 AA;

Query Match 50.8%; Score 185; DB 20; Length 185;  
Best Local Similarity 100.0%; Pred. No. 2.1e-179;  
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 12 MRYLLPSVVLGTAPRYVLAAGWWRLLSAFLPARFYQALDDRLXCYQSWLFEFFNYTG 71  
DB 1 mryllpsvvlgtapryvlawgwrllsalfparfyqalddrlxcyqswvlffenytg 60  
OY 72 VOILLGDLPRKKNENITTYLANHOSVDWIVADILAIRONALGHVRYVLEKGLKWLPLYGC 131  
DB 61 vqilllgdlprkknentitylanhosvdwivadilairgnalghvryvlekgllkwlpllygc 120  
OY 132 YFAOHGIIYVKRSKAKNEKEMRNKLSQSYDAGTGMTLVIFPECTRYNPEOTKVLASQAF 191  
DB 121 yfaohggiyvkrsaknekeemrnlqsyvdagtpmylvifpegtrynpeotkvlsasqaf 180  
OY 192 AAORG 196  
DB 181 aadrg 185

RESULT 15  
AAY36751

ID AAY36751 standard; Protein; 315 AA.  
XX  
XX AAY36751;  
XX  
XX 27-SEP-1999 (first entry)  
XX  
XX  
DE PGI splice variant allele protein sequence.  
XX  
XX PGI gene; biallelic marker; human; PSA; PGI-related biallelic marker;  
KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen.  
XX  
XX Homo sapiens.  
XX  
PN WO932644-A2.  
XX  
PD 01-JUL-1999.  
XX  
PF 22-DEC-1998; 98WO-IB02133.  
XX  
XX 09-SEP-1998; 98US-0099658.  
PR 22-DEC-1997; 97US-0996306.  
XX  
XX (GEST ) GENSET.  
XX  
PI Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;  
XX  
XX WPI: 1999-405178/34.  
DR N-PSDB; AA200977.  
XX  
XX Use of a prostate cancer associated gene and biallelic markers  
PT derived from it  
XX  
XX Claim 7; Page 253-254; 385pp; English.  
XX  
XX The invention relates to a mammalian PGI gene and protein, and a set of  
CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are  
CC used in a hybridisation assay, a sequencing assay, or in an  
CC allele-specific amplification assay for determining the identity of a  
CC nucleotide at a PGI-related biallelic marker. The methods can be used to  
CC detect and to assess the risk of developing cancer or prostate cancer.  
CC Early-stage diagnosis of prostate cancer relies on prostate specific  
CC antigen (PSA) dosage. However, the effectiveness of this is limited due  
CC to its inability to discriminate between malignant and non-malignant  
CC affections of the organ. A need exists for both a reliable diagnostic  
CC procedure which would enable early-stage diagnosis, and for preventative  
CC and curative treatments of the disease. The PGI gene can be used for  
CC detection of prostate cancer, and the risk of developing it in the  
CC future, and can also be used to determine therapies for the disease.  
XX  
XX Sequence 315 AA;

Query Match 50.8%; Score 185; DB 20; Length 315;  
Best Local Similarity 100.0%; Pred. No. 3.5e-179;  
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 12 MRYLLPSVVLGTAPRYVLAAGWWRLLSAFLPARFYQALDDRLXCYQSWLFEFFNYTG 71  
DB 1 mryllpsvvlgtapryvlawgwrllsalfparfyqalddrlxcyqswvlffenytg 60  
OY 72 VOILLGDLPRKKNENITTYLANHOSVDWIVADILAIRONALGHVRYVLEKGLKWLPLYGC 131  
DB 61 vqilllgdlprkknentitylanhosvdwivadilairgnalghvryvlekgllkwlpllygc 120  
OY 132 YFAOHGIIYVKRSKAKNEKEMRNKLSQSYDAGTGMTLVIFPECTRYNPEOTKVLASQAF 191  
DB 121 yfaohggiyvkrsaknekeemrnlqsyvdagtpmylvifpegtrynpeotkvlsasqaf 180  
OY 192 AAORG 196  
DB 181 aadrg 185

Search completed: August 28, 2002, 11:18:44  
Job time: 582 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:19:21 ; Search time 30.82 Seconds  
(without alignments)  
288.479 Million cell updates/sec

Title: US-09-853-526-5

Perfect score: 364  
Sequence: 1 MLSTLVHRYSMRYLLPSVY.....YVNTWYTGILLGLWYTKA 364

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 231628 seqs, 24425594 residues

Word size : 0

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

- Issued\_Patents\_AA:\*
- 1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*
  - 2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*
  - 3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*
  - 4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*
  - 5: /cgn2\_6/ptodata/2/1aa/PCtUS.COMB.pep:\*
  - 6: /cgn2\_6/ptodata/2/1aa/Backfillseq1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	364	100.0	364	2	US-08-996-306-5
2	364	100.0	364	4	US-09-338-907-5
3	364	100.0	364	4	US-09-218-207-5
4	353	97.0	353	2	US-08-996-306-4
5	353	97.0	353	4	US-09-338-907-4
6	353	97.0	353	4	US-09-218-207-4
7	230	63.2	291	4	US-09-338-907-127
8	230	63.2	291	4	US-09-218-207-127
9	228	62.6	228	4	US-09-338-907-70
10	228	62.6	228	4	US-09-218-207-70
11	199	54.7	261	4	US-09-338-907-128
12	199	54.7	261	4	US-09-218-207-128
13	185	50.8	185	4	US-09-338-907-136
14	185	50.8	185	4	US-09-218-207-136
15	185	50.8	185	4	US-09-338-907-134
16	185	50.8	185	4	US-09-338-907-134
17	184	50.5	300	4	US-09-338-907-135
18	184	50.5	300	4	US-09-218-207-135
19	154	42.3	182	4	US-09-338-907-133
20	154	42.3	182	4	US-09-218-207-133
21	116	31.9	238	4	US-09-338-907-126
22	116	31.9	238	4	US-09-218-207-126
23	85	23.4	97	4	US-09-338-907-132
24	85	23.4	97	4	US-09-218-207-132
25	62	17.0	66	4	US-09-338-907-131
26	62	17.0	66	4	US-09-218-207-131
27	62	17.0	68	4	US-09-338-907-130

28	62	17.0	68	4	US-09-218-207-130	Sequence 130, App
29	62	17.0	77	4	US-09-338-907-125	Sequence 125, App
30	62	17.0	77	4	US-09-218-207-125	Sequence 125, App
31	62	17.0	90	4	US-09-338-907-129	Sequence 129, App
32	62	17.0	90	4	US-09-218-207-129	Sequence 129, App
33	32	8.8	354	4	US-09-338-907-74	Sequence 74, App1
34	32	8.8	354	4	US-09-218-207-74	Sequence 74, App1
35	9	2.5	9	2	US-08-996-306-15	Sequence 15, App1
36	9	2.5	9	4	US-09-338-907-15	Sequence 15, App1
37	9	2.5	9	4	US-09-218-207-15	Sequence 15, App1
38	8	2.2	9	2	US-08-996-306-18	Sequence 18, App1
39	8	2.2	9	4	US-09-338-907-18	Sequence 18, App1
40	8	2.2	9	4	US-09-218-207-18	Sequence 18, App1
41	7	1.9	12	1	US-08-224-625-1	Sequence 1, App1
42	7	1.9	12	2	US-08-254-404-1	Sequence 1, App1
43	7	1.9	12	2	US-08-327-451E-1	Sequence 1, App1
44	7	1.9	12	2	US-08-458-109-1	Sequence 1, App1
45	7	1.9	12	3	US-08-231-196-1	Sequence 1, App1

ALIGNMENTS

RESULT 1  
US-08-996-306-5  
; Sequence 5, Application US/08996306  
; Patent No. 5945522  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Chumakov, Ilya  
; APPLICANT: Blumenfeld, Lydia  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: Prostate cancer gene  
; NUMBER OF SEQUENCES: 68  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Knobbe, Martens, Olson & Bear  
; STREET: 501 West Broadway  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92101-3505  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy Disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Win95  
; SOFTWARE: Word  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/996,306  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israelson, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: GENSET.018A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 364 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: SINGLE  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: PROTEIN  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: potential protein  
; LOCATION: 1..364  
; US-08-996-306-5

Query Match 100.0%; Score 364; DB 2; Length 364;

```
Best Local Similarity 100.0%; Pred. No. 0;
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLTSLVHTYSMRXYLLPSVLLGTAPTYYLAWGWRLLSAFLPARFYQALDRLCYVOS 60
Db 1 MLTSLVHTYSMRXYLLPSVLLGTAPTYYLAWGWRLLSAFLPARFYQALDRLCYVOS 60
Qy 61 MYLFFENYTGVOILLYGDLPRKNENIYLANHOSVDMIVADILAIRONALGHVRYVAK 120
Db 61 MYLFFENYTGVOILLYGDLPRKNENIYLANHOSVDMIVADILAIRONALGHVRYVAK 120
Qy 121 EGIKWLPLYGCTFAOHGGIYVKRSKAFNEKEMRNKLQSYVDAGTPMYLVIFFEGTRYNPE 180
Db 121 EGIKWLPLYGCTFAOHGGIYVKRSKAFNEKEMRNKLQSYVDAGTPMYLVIFFEGTRYNPE 180
Qy 181 QTKVLSASQAFAAQROGLAVLKHVLTPIKATHVAFDCMKNYLDAYDVYVYEGKDDGQ 240
Db 181 QTKVLSASQAFAAQROGLAVLKHVLTPIKATHVAFDCMKNYLDAYDVYVYEGKDDGQ 240
Qy 241 RRESPTMTEFLCKECPKIHIDRIDKDVPEBOEHMRMLHERFEIKDKMLIEFYESP 300
Db 241 RRESPTMTEFLCKECPKIHIDRIDKDVPEBOEHMRMLHERFEIKDKMLIEFYESP 300
Qy 301 PERRRKFPKGSVNSKLSIKTKLPMSLLSGLTAGMLMTDAGRKYVNTWYTGTLGCLMV 360
Db 301 PERRRKFPKGSVNSKLSIKTKLPMSLLSGLTAGMLMTDAGRKYVNTWYTGTLGCLMV 360
Qy 361 TTKA 364
Db 361 TTKA 364

RESULT 2
US-09-338-907-5
; Sequence 5, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CPICP
; CURRENT APPLICATION NUMBER: US/09/338,907
; CURRENT FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218,207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 5
; LENGTH: 364
; TYPE: prt
; ORGANISM: Homo sapiens
US-09-338-907-5

Query Match 100.0%; Score 364; DB 4; Length 364;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Db 121 EGIKWLPLYGCTFAOHGGIYVKRSKAFNEKEMRNKLQSYVDAGTPMYLVIFFEGTRYNPE 180
Qy 181 QTKVLSASQAFAAQROGLAVLKHVLTPIKATHVAFDCMKNYLDAYDVYVYEGKDDGQ 240
Db 181 QTKVLSASQAFAAQROGLAVLKHVLTPIKATHVAFDCMKNYLDAYDVYVYEGKDDGQ 240
Qy 241 RRESPTMTEFLCKECPKIHIDRIDKDVPEBOEHMRMLHERFEIKDKMLIEFYESP 300
Db 241 RRESPTMTEFLCKECPKIHIDRIDKDVPEBOEHMRMLHERFEIKDKMLIEFYESP 300
Qy 301 PERRRKFPKGSVNSKLSIKTKLPMSLLSGLTAGMLMTDAGRKYVNTWYTGTLGCLMV 360
Db 301 PERRRKFPKGSVNSKLSIKTKLPMSLLSGLTAGMLMTDAGRKYVNTWYTGTLGCLMV 360
Qy 361 TTKA 364
Db 361 TTKA 364

RESULT 3
US-09-218-207-5
; Sequence 5, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; FILE REFERENCE: GENSET.018CP1
; CURRENT APPLICATION NUMBER: US/09/218,207
; CURRENT FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 5
; LENGTH: 364
; TYPE: prt
; ORGANISM: Homo sapiens
US-09-218-207-5

Query Match 100.0%; Score 364; DB 4; Length 364;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



OY 361 TIKa 364  
DB 361 TIKa 364

RESULT 4  
US-08-996-306-4  
Sequence 4, Application US/08996306  
Patent No. 5945522  
GENERAL INFORMATION:  
APPLICANT: Cohen, Daniel  
APPLICANT: Chumakov, Ilya  
APPLICANT: Blumenfeld, Marta  
APPLICANT: Bouguenellet, Lydie  
TITLE OF INVENTION: Prostate cancer gene  
NUMBER OF SEQUENCES: 68  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobbe, Martens, Olson & Bear  
STREET: 501 West Broadway  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92101-3505  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Win95  
SOFTWARE: Word  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/996,306  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Israelsen, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: GENSET.018A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 235-8550  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 353 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: SINGLE  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PROTEIN  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 1..33  
IDENTIFICATION METHOD: Rao and Argos method  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 4..20  
IDENTIFICATION METHOD: Klein, Kanehisa and Delisi method  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 4..24  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komaromy  
IDENTIFICATION METHOD: and Wall method  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 12  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 50..70  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komaromy and Wall method  
FEATURE:  
NAME/KEY: potential N-glycosylation site  
LOCATION: 57

IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 76..96  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komaromy and Wall method  
FEATURE:  
NAME/KEY: potential Tyrosine kinase phosphorylation site  
LOCATION: 78  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Casein kinase II phosphorylation site  
LOCATION: 84  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Leucine zipper pattern  
LOCATION: 94..115  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 119  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 133  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Casein kinase II phosphorylation site  
LOCATION: 147  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 194  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Tyrosine kinase phosphorylation site  
LOCATION: 215  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Tyrosine sulfatation site  
LOCATION: 221  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential CAMP- and GMP-dependent protein kinase phosphorylation  
LOCATION: 233  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Casein kinase II phosphorylation site  
LOCATION: 235  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Protein kinase C phosphorylation site  
LOCATION: 306  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 310..330  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komaromy and Wall method  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 319  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential N-myristoylation site  
LOCATION: 323  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Amidation site  
LOCATION: 329  
IDENTIFICATION METHOD: prosite match  
FEATURE:  
NAME/KEY: potential Transmembrane helix  
LOCATION: 333..353  
IDENTIFICATION METHOD: Eisenberg, Schwarz, Komaromy and Wall method

```

;
; FEATURE:
; NAME/KEY: potential N-myristoylation site
; LOCATION: 341
; IDENTIFICATION METHOD: prosite match
; FEATURE:
; NAME/KEY: potential Protein kinase C phosphorylation site
; LOCATION: 350
; IDENTIFICATION METHOD: prosite match
; US-08-996-306-4

Query Match          97.0%: Score 353; DB 2; Length 353;
Best Local Similarity 100.0%: Pred. No. 0;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 MRYLLPSVVLGTAPTVYVLMGVRLSAFLPARFYQALDDRLCYVQSNVLEFFENYTG 71
    |||||
DB 1 MRYLLPSVVLGTAPTVYVLMGVRLSAFLPARFYQALDDRLCYVQSNVLEFFENYTG 60
    |||||
QY 72 VOILLVGDDLPKKNENIYYLANHOSYVDWIVADILAIQNALGHVRYVLAEGKWLPLYGC 131
    |||||
DB 61 VOILLVGDDLPKKNENIYYLANHOSYVDWIVADILAIQNALGHVRYVLAEGKWLPLYGC 120
    |||||
QY 132 YFQHGSIYKRSKFKENKEMRNKLSYVDAGTPMYLVIFPEGRVYNEQTKVLASQAF 191
    |||||
DB 121 YFQHGSIYKRSKFKENKEMRNKLSYVDAGTPMYLVIFPEGRVYNEQTKVLASQAF 180
    |||||
QY 192 AAGRGVLAVLKHVLTPIKATVAVFDCMKNYLDAIYDVTYVEGKDGGQRRSEPTMTTEFL 251
    |||||
DB 181 AAGRGVLAVLKHVLTPIKATVAVFDCMKNYLDAIYDVTYVEGKDGGQRRSEPTMTTEFL 240
    |||||
QY 252 CKKCPKTHIHIDRIKDVDEEEDHMRWLHERFEIKDKLIEFPSPDPERKRRFPGKS 311
    |||||
DB 241 CKKCPKTHIHIDRIKDVDEEEDHMRWLHERFEIKDKLIEFPSPDPERKRRFPGKS 300
    |||||
QY 312 VNSKLSIKTKLPMSLLISGLTAGMLTDAGRKIYVNTWITGTLGCLWVITKA 364
    |||||
DB 301 VNSKLSIKTKLPMSLLISGLTAGMLTDAGRKLIVNTWITGTLGCLWVITKA 353
    |||||

RESULT 5
US-09-338-907-4
; Sequence 4, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marica
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CPICP
; CURRENT APPLICATION NUMBER: US/09/338,907
; EARLIER FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218,207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 4
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: HELIX
; LOCATION: 1..33
; OTHER INFORMATION: Rao and Argos identification method, potential helix
; FEATURE:
; NAME/KEY: HELIX
; LOCATION: 4..20
; OTHER INFORMATION: Klein, Kanehisa and Delist identification method, potential]

;
; FEATURE:
; NAME/KEY: HELIX
; LOCATION: 4..24
; OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
; FEATURE:
; NAME/KEY: MYRISTATE
; LOCATION: 12..16
; OTHER INFORMATION: prosite match
; FEATURE:
; NAME/KEY: HELIX
; LOCATION: 50..70
; OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
; FEATURE:
; NAME/KEY: CARBOHYD
; LOCATION: 57..59
; OTHER INFORMATION: prosite match
; FEATURE:
; NAME/KEY: HELIX
; LOCATION: 76..96
; OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 78
; OTHER INFORMATION: potential Tyrosine kinase site, prosite match
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 84
; OTHER INFORMATION: potential caseline kinase II site, prosite match
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 94..115
; OTHER INFORMATION: potential Leucine zipper site, prosite match
; FEATURE:
; NAME/KEY: MYRISTATE
; LOCATION: 119..123
; OTHER INFORMATION: potential site, prosite match
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 133
; OTHER INFORMATION: potential protein kinase C, prosite match
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 147
; OTHER INFORMATION: potential caseline kinase II site, prosite match
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 194
; OTHER INFORMATION: potential protein kinase C, prosite match
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 215
; OTHER INFORMATION: potential Tyrosine kinase site, prosite match
; FEATURE:
; NAME/KEY: SULFATATION
; LOCATION: 221
; OTHER INFORMATION: prosite match
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 235
; OTHER INFORMATION: potential caseline kinase II site, prosite match
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 306
; OTHER INFORMATION: potential protein kinase C, prosite match
; FEATURE:
; NAME/KEY: HELIX
; LOCATION: 310..330
; OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
; FEATURE:
```

```

? NAME/KEY: MYRISTATE
? LOCATION: 319..323
? OTHER INFORMATION: Prosite match
? FEATURE:
? NAME/KEY: MYRISTATE
? LOCATION: 323..327
? OTHER INFORMATION: Prosite match
? FEATURE:
? NAME/KEY: AMIDATION
? LOCATION: 329
? OTHER INFORMATION: Prosite match
? FEATURE:
? NAME/KEY: HELIX
? LOCATION: 333..353
? OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method.
? FEATURE:
? NAME/KEY: MYRISTATE
? LOCATION: 341..345
? OTHER INFORMATION: Prosite match
? FEATURE:
? NAME/KEY: PHOSPHORYLATION
? LOCATION: 350
? OTHER INFORMATION: potential protein kinase C, Prosite match
? PS-09-338-907-4

```

```

RESULT 6
US-09-218-207-4
: Sequence 4, Application US/09218207
: Patent No. 6346381
:
: GENERAL INFORMATION:
:
: APPLICANT: Cohen, Daniel
: APPLICANT: Blumenfeld, Marta
: APPLICANT: Ilya, Chumakov
: APPLICANT: Bougueleret, Lydie
: TITLE OF INVENTION: Prostate cancer gene
: FILE REFERENCE: GENSET.018CP1
: CURRENT APPLICATION NUMBER: US/09/218,207
: CURRENT FILING DATE: 1998-12-22
: EARLIER APPLICATION NUMBER: 08/996,306
: EARLIER FILING DATE: 1997-12-22
: EARLIER APPLICATION NUMBER: 60/099,658
: EARLIER FILING DATE: 1996-09-09
: NUMBER OF SEQ ID NOS: 578

```

```

1 SOFTWARE: Patent.pm
2 SEQ ID NO 4
3 LENGTH: 353
4 TYPE: PRT
5 ORGANISM: Homo sapiens
6 FEATURE:
7 NAME/KEY: HELIX
8 LOCATION: 1..33
9 OTHER INFORMATION: Rao and Argos identification method, potential helix
10 FEATURE:
11 NAME/KEY: HELIX
12 LOCATION: 4..20
13 OTHER INFORMATION: Klein, Kanehisa and Delisi identification method, potent
14 FEATURE:
15 NAME/KEY: HELIX
16 LOCATION: 4..24
17 OTHER INFORMATION: Eisenberg, Schwarz, Komarony, wall identification method,
18 FEATURE:
19 NAME/KEY: MYRISTATE
20 LOCATION: 12..16
21 OTHER INFORMATION: Prosite match
22 FEATURE:
23 NAME/KEY: HELIX
24 LOCATION: 50..70
25 OTHER INFORMATION: Eisenberg, Schwarz, Komarony, wall identification method,
26 FEATURE:
27 NAME/KEY: CARBOHYD
28 LOCATION: 57..59
29 OTHER INFORMATION: Prosite match
30 FEATURE:
31 NAME/KEY: HELIX
32 LOCATION: 76..96
33 OTHER INFORMATION: Eisenberg, Schwarz, Komarony, wall identification method,
34 FEATURE:
35 NAME/KEY: PHOSPHORYLATION
36 LOCATION: 78
37 OTHER INFORMATION: potential Tyrosine kinase site, Prosite match
38 FEATURE:
39 NAME/KEY: PHOSPHORYLATION
40 LOCATION: 84
41 OTHER INFORMATION: potential caseine kinase II site, Prosite match
42 FEATURE:
43 NAME/KEY: SITE
44 LOCATION: 94..115
45 OTHER INFORMATION: potential leucine zipper site, Prosite match
46 FEATURE:
47 NAME/KEY: MYRISTATE
48 LOCATION: 119..123
49 OTHER INFORMATION: potential site, Prosite match
50 FEATURE:
51 NAME/KEY: PHOSPHORYLATION
52 LOCATION: 133
53 OTHER INFORMATION: potential protein kinase C, Prosite match
54 FEATURE:
55 NAME/KEY: PHOSPHORYLATION
56 LOCATION: 147
57 OTHER INFORMATION: potential caseine kinase II site, Prosite match
58 FEATURE:
59 NAME/KEY: PHOSPHORYLATION
60 LOCATION: 194
61 OTHER INFORMATION: potential protein kinase C, Prosite match
62 FEATURE:
63 NAME/KEY: PHOSPHORYLATION
64 LOCATION: 215
65 OTHER INFORMATION: potential Tyrosine kinase site, Prosite match
66 FEATURE:
67 NAME/KEY: SULFATATION
68 LOCATION: 221
69 OTHER INFORMATION: Prosite match
70 FEATURE:
71 NAME/KEY: PHOSPHORYLATION
72 LOCATION: 233
73 OTHER INFORMATION: potential cAMP and cGMP dependant protein kinase site,

```

```

; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 235
; OTHER INFORMATION: potential caseine kinase II site, Prosite match
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 306
; OTHER INFORMATION: potential protein kinase C, Prosite match
; FEATURE:
; NAME/KEY: HELIX
; LOCATION: 310..330
; OTHER INFORMATION: Eisenberg, Schwarz, Komaromy, Wall identification method, F
; FEATURE:
; NAME/KEY: MYRISTATE
; LOCATION: 319..323
; OTHER INFORMATION: Prosite match
; FEATURE:
; NAME/KEY: MYRISTATE
; LOCATION: 323..327
; OTHER INFORMATION: Prosite match
; FEATURE:
; NAME/KEY: AMIDATION
; LOCATION: 329
; OTHER INFORMATION: Prosite match
; FEATURE:
; NAME/KEY: HELIX
; LOCATION: 333..353
; OTHER INFORMATION: Eisenberg, Schwarz, Komaromy, Wall identification method, F
; FEATURE:
; NAME/KEY: MYRISTATE
; LOCATION: 341..345
; OTHER INFORMATION: Prosite match
; FEATURE:
; NAME/KEY: PHOSPHORYLATION
; LOCATION: 350
; OTHER INFORMATION: potential protein kinase C, Prosite match
; US-09-218-207-4
```

```

Query Match          97.0%; Score 353; DB 4; Length 353;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 12 MRYLPSVVLGTAPYVLAAGVWLLSAPLPARFYQALDDRLICYQSNLFEFENYTG 71
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1 MRYLPSVVLGTAPYVLAAGVWLLSAPLPARFYQALDDRLICYQSNLFEFENYTG 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 72 VOILLVGDLPKKNKENTIIYLAHOSYVDWIVADILAIROMALGHVRYVLEKGLWPLVYGC 131
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 61 VOILLVGDLPKKNKENTIIYLAHOSYVDWIVADILAIROMALGHVRYVLEKGLWPLVYGC 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 132 YFAOHGGIYVKSASAKNEKEMRNKLOSVDAGTPMYLVIFPESTRYNPEQTKVLSASQAF 191
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 121 YFAOHGGIYVKSASAKNEKEMRNKLOSVDAGTPMYLVIFPESTRYNPEQTKVLSASQAF 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 192 AAQGLAVLKHLVTPRIKATHTVAFDCKMKNYLDIYDVTVVYEGKDDGGRRESPTWTEFL 251
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 181 AAQGLAVLKHLVTPRIKATHTVAFDCKMKNYLDIYDVTVVYEGKDDGGRRESPTWTEFL 240
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 252 CEECPRIHIDRIKDDVPEDEHMRWLHERFEIKDKMLIEFESPPDERRRKRPGRS 311
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 241 CEECPRIHIDRIKDDVPEDEHMRWLHERFEIKDKMLIEFESPPDERRRKRPGRS 300
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 312 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLKYVNTWITYGTLGLWVTIKR 364
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 301 VNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLKYVNTWITYGTLGLWVTIKR 353
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```

RESULT 7
US-09-338-907-127
; Sequence 127, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
```

```

; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CPLCP
; CURRENT APPLICATION NUMBER: US/09/338,907
; EARLIER FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218,207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 127
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 98..103
; OTHER INFORMATION: Box II
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 149..157
; OTHER INFORMATION: Box III
; US-09-338-907-127
```

```

Query Match          63.2%; Score 230; DB 4; Length 291;
Best Local Similarity 100.0%; Pred. No. 2.9e-223;
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 135 QHGGIYVKSASAKNEKEMRNKLOSVDAGTPMYLVIFPESTRYNPEQTKVLSASQAF 194
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 62 QHGGIYVKSASAKNEKEMRNKLOSVDAGTPMYLVIFPESTRYNPEQTKVLSASQAF 121
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 195 RGLAVLKHLVTPRIKATHTVAFDCKMKNYLDIYDVTVVYEGKDDGGRRESPTWTEFLCKE 254
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 122 RGLAVLKHLVTPRIKATHTVAFDCKMKNYLDIYDVTVVYEGKDDGGRRESPTWTEFLCKE 181
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 255 CPEKIHIDRIKDDVPEDEHMRWLHERFEIKDKMLIEFESPPDERRRKRPGRSVNS 314
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 182 CPEKIHIDRIKDDVPEDEHMRWLHERFEIKDKMLIEFESPPDERRRKRPGRSVNS 241
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 315 KLSIKKTLPSMLISGLTAGMLMTDAGRKLKYVNTWITYGTLGLWVTIKR 364
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 242 KLSIKKTLPSMLISGLTAGMLMTDAGRKLKYVNTWITYGTLGLWVTIKR 291
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```

RESULT 8
US-09-218-207-127
; Sequence 127, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; FILE REFERENCE: GENSET.018CPL
; CURRENT APPLICATION NUMBER: US/09/218,207
; EARLIER FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 127
; LENGTH: 291
```

```

? TYPE: PRT
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: SITE
? LOCATION: 98..103
? OTHER INFORMATION: Box
? FEATURE:
? NAME/KEY: SITE
? LOCATION: 149..157
? OTHER INFORMATION: Box
US-09-218-007-127

```

Query Match	63.2%	Score 230;	DB 4;	Length 291;
Best Local Similarity	100.0%;	Pred. No. 2.9e+223;		
Matches 230;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	135	OHGGIYKRSKAFNEKEEMRNKLOSYVAGPMPYLVIPPEGRNRPBETKLSASQAAQ	154
Db	62	OHGGIYKRSKAFNEKEEMRNKLOSYVAGPMPYLVIPPEGRNRPBETKLSASQAAQ	122
QY	195	RGIAVLKHLVLPRIKATHVAAPDCAKNLTDAIYDTVVYEEGKDQGGQRESPTMEFFCKE	254
Db	122	RGIAVLKHLVLPRIKATHVAAPDCAKNLTDAIYDTVVYEEGKDQGGQRESPTMEFFCKE	161
QY	255	CPKTHIHIDRIDKKDVPEDEEHMRMLHERFEIKDKMLIEFYESPDPERKKRPPGKSVNS	314
Db	182	CPKTHIHIDRIDKKDVPEDEEHMRMLHERFEIKDKMLIEFYESPDPERKKRPPGKSVNS	241
QY	315	KLSIKTKLPMSMLISGLTAGMLMDAGRKLVNWNVYIGTLGLCLWATIKAKA	364
Db	242	KLSIKTKLPMSMLISGLTAGMLMDAGRKLVNWNVYIGTLGLCLWATIKAKA	291

```

RESULT 9
US-09-338-907-70
Sequence 70, Application US/09338907
Patent No. 6265346
GENERAL INFORMATION:
APPLICANT: Cohen, Daniel
APPLICANT: Blumenfeld, Marta
APPLICANT: Ilyva, Chumakov
APPLICANT: Bougueleret, Lydie
TITLE OF INVENTION: PROSTATE CANCER GENE
FILE REFERENCE: GENSET.18CPLC
CURRENT APPLICATION NUMBER: US/09/338, 907
CURRENT FILING DATE: 1999-06-23
EARLIER APPLICATION NUMBER: 08/996, 306
EARLIER FILING DATE: 1997-12-22
EARLIER APPLICATION NUMBER: 60/059, 658
EARLIER FILING DATE: 1998-09-09
EARLIER APPLICATION NUMBER: 09/218, 207
EARLIER FILING DATE: 1998-12-22
NUMBER OF SEQ ID NOS: 578
SOFTWARE: Patent.pm
SEQ ID NO 70
LENGTH: 228
TYPE: PRT
ORGANISM: Homo sapiens
US-09-338-907-70

```

Query Match	62.6%	Score 228;	DB 4;	Length 228;
Best Local Similarity	100.0%	Pred. No. 2,4e-221;		
Matches 228; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0

0y	12	MRVLLPSVLLGTPPTVYIAGVNRLLSAPFPAEFOALDRICVQSNVLPFEENYTG	71
Db	1	MRVLLPSVLLGTPPTVYIAGVNRLLSAPFPAEFOALDRICVQSNVLPFEENYTG	60
0y	72	VQILLGDLPRNKENIIYLANHOSFVDMIVADIIAIRONALGHRVYVLEKGLKMLPYGC	131
Db	61	VQILLGDLPRNKENIIYLANHOSFVDMIVADIIAIRONALGHRVYVLEKGLKMLPYGC	120

Qy	132	YFHHGGIYYVRSKAFENKEKMRNKLQSVADGPMYIVIEPEGRIYRNPEDQTKYLSAQA	129
Db	121	YFAHGGIYYVRSKAFENKEKMRNKLQSVADGPMYIVIEPEGRIYRNPEDQTKYLSAQA	180
Qy	132	AAQRCGLAVLKHVLPRIKATHVAEDCKMNYLDALYDVTYYEGKDDGG	239
Db	181	AAQRCGLAVLKHVLPRIKATHVAEDCKMNYLDALYDVTYYEGKDDGG	228

```

10
US-09-218-207-70
; Sequence 70, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer ge
; FILE REFERENCE: GENSET. 018CPI
; CURRENT APPLICATION NUMBER: US/09/218,
; EARLIER FILING DATE: 1998-12-22
; EARLY APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 70
; LENGTH: 228
; TYPE: prt
; ORGANISM: Homo sapiens
US-09-218-207-70

```

	Query Match	Similarity	62.6%	Score 228	DB 4	Length 228
	Best Local	Similarity	100.0%	Pred. No. 2.4e-221		
	Matches	228	Conservative	0	Mismatches	0
					Indels	0
					Gaps	0
QY	12	MRLLPSVLLGTA	PTTYLVAMGWMRLSAFLPAREYQALDRLKLCVYOSMVLFFENN	YTG	71	
Db	1	MRLLPSVLLGTA	PTTYLVAMGWMRLSAFLPAREYQALDRLKLCVYOSMVLFFENN	YTG	60	
QY	72	VOIILYGDLPKK	ENITYLANHOSVNDVVAIILAIROMALGHVRYVLRKEGKMLPIYGC	131		
Db	61	VOIILYGDLPKK	ENITYLANHOSVNDVVAIILAIROMALGHVRYVLRKEGKMLPIYGC	120		
QY	133	YFAQHGGIYVKS	AKAFENEKEMRNKLQSYVDAC	TPMYLFFPEGRTRNPEQRTVLSASQAF	191	
Db	121	YFAQHGGIYVKS	AKAFENEKEMRNKLQSYVDAC	TPMYLFFPEGRTRNPEQRTVLSASQAF	180	
QY	132	AAORGLAVLKH	VLTPRIKATHVAFDCMKNYLDAITDYVYVYSGKDDG	239		
Db	181	AAORGLAVLKH	VLTPRIKATHVAFDCMKNYLDAITDYVYVYSGKDDG	228		

RESULT 11  
US-09-338-907-128

RESULT 11  
 US-09-338-907-128  
 : Sequence 128, Application US/09338907  
 : Patent No. 6265546  
 :  
 : GENERAL INFORMATION:  
 :  
 : APPLICANT: Cohen, Daniel  
 : APPLICANT: Blumenfeld, Marta  
 : APPLICANT: Ilya, Chumakov  
 : APPLICANT: Bougueleret, Lydie  
 :  
 : TITLE OF INVENTION: PROSTATE CANCER GENE  
 : FILE REFERENCE: GENSET.18CP  
 : CURRENT APPLICATION NUMBER: US/09/338, 907  
 : CURRENT FILING DATE: 1999-06-23  
 : EARLIER APPLICATION NUMBER: 08/996, 306  
 : EARLIER FILING DATE: 1997-12-22  
 : EARLIER APPLICATION NUMBER: 60/099, 658

```

; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218, 207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 128
; LENGTH: 261
; TYPE: PRF
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 68..73
; OTHER INFORMATION: Box II
; NAME/KEY: SITE
; LOCATION: 119..127
; OTHER INFORMATION: Box III
US-09-338-907-128
```

```

Query Match          54.7%; Score 199; DB 4; Length 261;
Best Local Similarity 100.0%; Pred. No. 3.9e-192;
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 166 MYLVIEPEGTRYNPEQTKVLSASQAFQAQGLAVLKHVLTLPRIKATHVAFDCMKNYLDAL 225
; |||||||
DB 63 MYLVIEPEGTRYNPEQTKVLSASQAFQAQGLAVLKHVLTLPRIKATHVAFDCMKNYLDAL 122
; |||||||
QY 226 YDVTVYVEGDDGGORRESPTMTPEFLCKECPKHHIDRDKDVPEDECHMRMLHERF 285
; |||||||
DB 123 YDVTVYVEGDDGGORRESPTMTPEFLCKECPKHHIDRDKDVPEDECHMRMLHERF 182
; |||||||
QY 286 EIKDKMLIEFESPDPERRRRPFKGSVNSKLSIKTLPMSLISGLTAGMLMTDAGRKL 345
; |||||||
DB 183 EIKDKMLIEFESPDPERRRRPFKGSVNSKLSIKTLPMSLISGLTAGMLMTDAGRKL 242
; |||||||
QY 346 VNTWYGTLLGCLMWTIKA 364
; |||||||
DB 243 VNTWYGTLLGCLMWTIKA 261
; |||||||
```

```

RESULT 12
US-09-218-207-128
; Sequence 128, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marita
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; FILE REFERENCE: GENSET.018CPI
; CURRENT APPLICATION NUMBER: US/09/218, 207
; EARLIER FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 08/996, 306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099, 658
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 128
; LENGTH: 261
; TYPE: PRF
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 68..73
; OTHER INFORMATION: Box II
; NAME/KEY: SITE
; LOCATION: 119..127
; OTHER INFORMATION: Box III
US-09-218-207-128
```

```

Query Match          54.7%; Score 199; DB 4; Length 261;
Best Local Similarity 100.0%; Pred. No. 3.9e-192;
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 MYLVIEPEGTRYNPEQTKVLSASQAFQAQGLAVLKHVLTLPRIKATHVAFDCMKNYLDAL 225
; |||||||
DB 63 MYLVIEPEGTRYNPEQTKVLSASQAFQAQGLAVLKHVLTLPRIKATHVAFDCMKNYLDAL 122
; |||||||
QY 226 YDVTVYVEGDDGGORRESPTMTPEFLCKECPKHHIDRDKDVPEDECHMRMLHERF 285
; |||||||
DB 123 YDVTVYVEGDDGGORRESPTMTPEFLCKECPKHHIDRDKDVPEDECHMRMLHERF 182
; |||||||
QY 286 EIKDKMLIEFESPDPERRRRPFKGSVNSKLSIKTLPMSLISGLTAGMLMTDAGRKL 345
; |||||||
DB 183 EIKDKMLIEFESPDPERRRRPFKGSVNSKLSIKTLPMSLISGLTAGMLMTDAGRKL 242
; |||||||
QY 346 VNTWYGTLLGCLMWTIKA 364
; |||||||
DB 243 VNTWYGTLLGCLMWTIKA 261
; |||||||
```

```

RESULT 13
US-09-338-907-136
; Sequence 136, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marita
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CPI
; CURRENT APPLICATION NUMBER: US/09/338, 907
; EARLIER FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996, 306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099, 658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218, 207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 136
; LENGTH: 185
; TYPE: PRF
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 81..83
; OTHER INFORMATION: Box I
; NAME/KEY: SITE
; LOCATION: 160..165
; OTHER INFORMATION: Box II
US-09-338-907-136
```

```

Query Match          50.8%; Score 185; DB 4; Length 185;
Best Local Similarity 100.0%; Pred. No. 3.3e-178;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 MYLPSVVLGTAPTYVILAMGVRLLSAFLPARFYQALDDRLCYVYQSMVLEFFENYTG 71
; |||||||
DB 1 MYLPSVVLGTAPTYVILAMGVRLLSAFLPARFYQALDDRLCYVYQSMVLEFFENYTG 60
; |||||||
QY 72 VOILLYGDLPRKKNENITILANHOSTYDWMIVADILAIQNALGHRVYLKEGLKMLPLYGC 131
; |||||||
DB 61 VOILLYGDLPRKKNENITILANHOSTYDWMIVADILAIQNALGHRVYLKEGLKMLPLYGC 120
; |||||||
QY 132 YFAOHGIGIVKRSARKNEKEMRKLOSVDAGPMVLYFPGSTRYNPEQTKVLSASQAF 191
; |||||||
```

```
Db 121 YFAOHGGIYKRSKAFNEKEMRNKLQSYVDAGT PMLVIFPECTRYNPQTKVLSASQAF 180
QY 192 AAORG 196
    |||||
Db 181 AAORG 185

RESULT 14
US-09-218-207-136
; Sequence 136, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; FILE REFERENCE: GENSET.018CP1
; CURRENT APPLICATION NUMBER: US/09/218,207
; CURRENT FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 136
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 81..83
; OTHER INFORMATION: Box I
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 160..165
; OTHER INFORMATION: Box II
US-09-218-207-136

Query Match          50.8%; Score 185; DB 4; Length 185;
Best Local Similarity 100.0%; Pred. No. 3,3e-178;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 MRYLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQSNVLEFFENYTG 71
    |||||
Db 1 MRYLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQSNVLEFFENYTG 60
QY 72 VOILLYGDLPKNKENIITYLANHOSTVDWIVADILAIRONALGHVRYVLEKGLKMLPLLYGC 131
    |||||
Db 61 VOILLYGDLPKNKENIITYLANHOSTVDWIVADILAIRONALGHVRYVLEKGLKMLPLLYGC 120
QY 132 YFAOHGGIYKRSKAFNEKEMRNKLQSYVDAGT PMLVIFPECTRYNPQTKVLSASQAF 191
    |||||
Db 121 YFAOHGGIYKRSKAFNEKEMRNKLQSYVDAGT PMLVIFPECTRYNPQTKVLSASQAF 180
QY 192 AAORG 196
    |||||
Db 181 AAORG 185

RESULT 15
US-09-338-907-134
; Sequence 134, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CP1CP
```

```
; CURRENT APPLICATION NUMBER: US/09/338,907
; CURRENT FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218,207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 134
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 81..83
; OTHER INFORMATION: Box I
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 160..165
; OTHER INFORMATION: Box II
US-09-338-907-134

Query Match          50.8%; Score 185; DB 4; Length 315;
Best Local Similarity 100.0%; Pred. No. 5,5e-178;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 MRYLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQSNVLEFFENYTG 71
    |||||
Db 1 MRYLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQSNVLEFFENYTG 60
QY 72 VOILLYGDLPKNKENIITYLANHOSTVDWIVADILAIRONALGHVRYVLEKGLKMLPLLYGC 131
    |||||
Db 61 VOILLYGDLPKNKENIITYLANHOSTVDWIVADILAIRONALGHVRYVLEKGLKMLPLLYGC 120
QY 132 YFAOHGGIYKRSKAFNEKEMRNKLQSYVDAGT PMLVIFPECTRYNPQTKVLSASQAF 191
    |||||
Db 121 YFAOHGGIYKRSKAFNEKEMRNKLQSYVDAGT PMLVIFPECTRYNPQTKVLSASQAF 180
QY 192 AAORG 196
    |||||
Db 181 AAORG 185
```

Search completed: August 28, 2002, 11:19:21  
Job time: 304 sec

1.



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2002, 11:25:19 ; Search time 305.3 Seconds  
(without alignments)  
419.656 Million cell updates/sec

Title: US-09-853-526-5

Perfect score: 364  
Sequence: 1 MLSTLVHTYSMRYLPSVY.....YVNTWLYGFLGLGLMTWTKA 364

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 3502263 seqs, 351980561 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main:\*

- 1: /cgn2\_6/ptodata/2/paa/PCYUS.COMB.pep:\*
- 2: /cgn2\_6/ptodata/2/paa/US06.COMB.pep:\*
- 3: /cgn2\_6/ptodata/2/paa/US07.COMB.pep:\*
- 4: /cgn2\_6/ptodata/2/paa/US081.COMB.pep:\*
- 5: /cgn2\_6/ptodata/2/paa/US081.COMB.pep:\*
- 6: /cgn2\_6/ptodata/2/paa/US083.COMB.pep:\*
- 7: /cgn2\_6/ptodata/2/paa/US083.COMB.pep:\*
- 8: /cgn2\_6/ptodata/2/paa/US084.COMB.pep:\*
- 9: /cgn2\_6/ptodata/2/paa/US085.COMB.pep:\*
- 10: /cgn2\_6/ptodata/2/paa/US086.COMB.pep:\*
- 11: /cgn2\_6/ptodata/2/paa/US087.COMB.pep:\*
- 12: /cgn2\_6/ptodata/2/paa/US088.COMB.pep:\*
- 13: /cgn2\_6/ptodata/2/paa/US089.COMB.pep:\*
- 14: /cgn2\_6/ptodata/2/paa/US090.COMB.pep:\*
- 15: /cgn2\_6/ptodata/2/paa/US091.COMB.pep:\*
- 16: /cgn2\_6/ptodata/2/paa/US092.COMB.pep:\*
- 17: /cgn2\_6/ptodata/2/paa/US093.COMB.pep:\*
- 18: /cgn2\_6/ptodata/2/paa/US094.COMB.pep:\*
- 19: /cgn2\_6/ptodata/2/paa/US095.COMB.pep:\*
- 20: /cgn2\_6/ptodata/2/paa/US096.COMB.pep:\*
- 21: /cgn2\_6/ptodata/2/paa/US097.COMB.pep:\*
- 22: /cgn2\_6/ptodata/2/paa/US098.COMB.pep:\*
- 23: /cgn2\_6/ptodata/2/paa/US099.COMB.pep:\*
- 24: /cgn2\_6/ptodata/2/paa/US100.COMB.pep:\*
- 25: /cgn2\_6/ptodata/2/paa/US101.COMB.pep:\*
- 26: /cgn2\_6/ptodata/2/paa/US60.COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	364	100.0	364	18	US-09-436-919-1
2	364	100.0	364	22	US-09-817-910-7
3	364	100.0	364	22	US-09-853-526-5
4	364	100.0	364	23	US-09-901-484A-5
5	364	100.0	364	26	US-60-099-658-5
6	353	97.0	353	1	PCT-US01-01431-59
7	353	97.0	353	1	PCT-US01-01431-79

8	353	97.0	353	1	PCT-US01-11988-1474	Sequence 1474, Ap
9	353	97.0	353	22	US-09-833-245-1475	Sequence 1475, Ap
10	353	97.0	353	22	US-09-833-245-1475	Sequence 1474, Ap
11	353	97.0	353	22	US-09-833-245-1475	Sequence 1475, Ap
12	353	97.0	353	22	US-09-833-245-1475	Sequence 4, Appl1
13	353	97.0	353	23	US-09-901-484A-4	Sequence 4, Appl1
14	353	97.0	353	23	US-09-915-582-59	Sequence 59, Appl1
15	353	97.0	353	23	US-09-915-582-79	Sequence 79, Appl1
16	353	97.0	353	26	US-60-099-658-4	Sequence 4, Appl1
17	290	79.7	353	18	US-09-488-725A-2736	Sequence 2736, Ap
18	269	73.9	269	1	PCT-US01-01327-47	Sequence 47, Appl1
19	263	72.3	372	18	US-09-488-725A-6308	Sequence 6308, Ap
20	248	68.1	450	26	US-60-245-222-110	Sequence 110, App
21	230	63.2	291	22	US-09-853-526-127	Sequence 127, App
22	228	62.6	228	22	US-09-901-484A-127	Sequence 70, Appl1
23	228	62.6	228	22	US-09-853-526-136	Sequence 127, App
24	228	62.6	228	23	US-09-901-484A-70	Sequence 70, Appl1
25	228	62.6	228	26	US-60-099-658-70	Sequence 70, Appl1
26	199	54.7	261	22	US-09-853-526-128	Sequence 128, App
27	199	54.7	261	23	US-09-901-484A-128	Sequence 128, App
28	185	50.8	185	22	US-09-853-526-136	Sequence 136, App
29	185	50.8	185	23	US-09-901-484A-136	Sequence 136, App
30	185	50.8	315	22	US-09-853-526-134	Sequence 134, App
31	185	50.8	315	23	US-09-901-484A-134	Sequence 134, App
32	184	50.5	300	22	US-09-853-526-135	Sequence 135, App
33	184	50.5	300	23	US-09-901-484A-135	Sequence 135, App
34	154	42.3	182	22	US-09-853-526-133	Sequence 133, App
35	154	42.3	182	23	US-09-901-484A-133	Sequence 133, App
36	153	42.0	176	18	US-09-436-919-5	Sequence 5, Appl1
37	151	41.5	1032	26	US-60-212-413-309	Sequence 309, App
38	151	41.5	1032	26	US-60-229-518-248	Sequence 248, App
39	116	31.9	238	22	US-09-853-526-126	Sequence 126, App
40	116	31.9	238	23	US-09-901-484A-126	Sequence 126, App
41	85	23.4	97	22	US-09-853-526-132	Sequence 132, App
42	85	23.4	97	23	US-09-901-484A-132	Sequence 132, App
43	81	22.3	92	17	US-09-316-123-124	Sequence 124, App
44	81	22.3	92	22	US-09-827-244-124	Sequence 124, App
45	70	19.2	257	1	PCT-US01-08631-31719	Sequence 31719, A

ALIGNMENTS

RESULT 1  
US-09-436-919-1  
; Sequence 1, Application US/09436919A  
; GENERAL INFORMATION:  
; APPLICANT: Leung, David W  
; TITLE OF INVENTION: Human Lysophosphatidic Acid Acyltransferase-epsilon  
; FILE REFERENCE: 1801B  
; CURRENT APPLICATION NUMBER: US/09/436, 919A  
; CURRENT FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 364  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: huPAREpsilon1  
US-09-436-919-1

Query Match	100.0%	Score 364;	DB 18;	Length 364;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 364;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MLSTLVHTYSMRYLPSVVLGTAFTYVLAMGVMLLSAFLPARFYQALDRLCYVYOS	60	
DB	1	MLSTLVHTYSMRYLPSVVLGTAFTYVLAMGVMLLSAFLPARFYQALDRLCYVYOS	60	
QY	61	MYLFFENTYGVQILLYGDLPRKNENIIVLANHSTVDIVADIIARONALGHVRYVLR	120	

Db 61 MVLFFENNTGVOILLXGDLPPKKNENIYLANHSTVDMIVADILAIRONALGHVRYVLK 120  
QY 121 EGLKWLPLYGCVFAOHGIIYVVKRSKAFNEKEMRNKLQSYVDAGTPMYLVIFPEGTRYNPE 180  
Db 121 EGLKWLPLYGCVFAOHGIIYVVKRSKAFNEKEMRNKLQSYVDAGTPMYLVIFPEGTRYNPE 180  
QY 181 QTKVLSASQAFAAORGLAVLKHVLTFRKATNHAFCMKNYLDAIYDVTVYVEGKDDGQ 240  
Db 181 QTKVLSASQAFAAORGLAVLKHVLTFRKATNHAFCMKNYLDAIYDVTVYVEGKDDGQ 240  
QY 241 RRESPTMTFELCKECPKIHIDRIDKDVPEOEHRMRHLHERFEIKDKMLIEFYESPD 300  
Db 241 RRESPTMTFELCKECPKIHIDRIDKDVPEOEHRMRHLHERFEIKDKMLIEFYESPD 300  
QY 301 PERRRRPCKSVNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTWYTWIYGTLLGCLMW 360  
Db 301 PERRRRPCKSVNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTWYTWIYGTLLGCLMW 360  
QY 361 TTKA 364  
Db 361 TTKA 364

RESULT 2  
US-09-817-910-7  
; Sequence 7, Application US/09817910  
; GENERAL INFORMATION:  
; APPLICANT: Meyers, Rachel A.  
; APPLICANT: Macbeth, Kyle J.  
; APPLICANT: Williamson, Mark  
; APPLICANT: Rudolph-Owen, Laura A.  
; APPLICANT: Tsai, Fong-Ying  
; TITLE OF INVENTION: 46743 and 27417 NOVEL HUMAN  
; FILE REFERENCE: 10448-032001  
; CURRENT APPLICATION NUMBER: US/09/817,910  
; CURRENT FILING DATE: 2001-03-26  
; PRIOR APPLICATION NUMBER: US 60/192,092  
; PRIOR FILING DATE: 2000-03-24  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 364  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-817-910-7

Query Match 100.0%; Score 364; DB 22; Length 364;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLISLVLHHTYSMRYLLPSVVLGTAPTYVLAWGVWRLLSAFUPARYOALDRLXYCVOS 60  
Db 1 MLISLVLHHTYSMRYLLPSVVLGTAPTYVLAWGVWRLLSAFUPARYOALDRLXYCVOS 60  
QY 61 MVLFFENNTGVOILLXGDLPPKKNENIYLANHSTVDMIVADILAIRONALGHVRYVLK 120  
Db 61 MVLFFENNTGVOILLXGDLPPKKNENIYLANHSTVDMIVADILAIRONALGHVRYVLK 120  
QY 121 EGLKWLPLYGCVFAOHGIIYVVKRSKAFNEKEMRNKLQSYVDAGTPMYLVIFPEGTRYNPE 180  
Db 121 EGLKWLPLYGCVFAOHGIIYVVKRSKAFNEKEMRNKLQSYVDAGTPMYLVIFPEGTRYNPE 180  
QY 181 QTKVLSASQAFAAORGLAVLKHVLTFRKATNHAFCMKNYLDAIYDVTVYVEGKDDGQ 240  
Db 181 QTKVLSASQAFAAORGLAVLKHVLTFRKATNHAFCMKNYLDAIYDVTVYVEGKDDGQ 240  
QY 241 RRESPTMTFELCKECPKIHIDRIDKDVPEOEHRMRHLHERFEIKDKMLIEFYESPD 300  
Db 241 RRESPTMTFELCKECPKIHIDRIDKDVPEOEHRMRHLHERFEIKDKMLIEFYESPD 300  
QY 301 PERRRRPCKSVNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTWYTWIYGTLLGCLMW 360  
Db 301 PERRRRPCKSVNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTWYTWIYGTLLGCLMW 360

Db 301 PERRRRPCKSVNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTWYTWIYGTLLGCLMW 360  
QY 361 TTKA 364  
Db 361 TTKA 364

RESULT 3  
US-09-853-526-5  
; Sequence 5, Application US/09853526  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Ilya, Chumakov  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: PROSTATE CANCER GENE  
; FILE REFERENCE: GENSET.18CP1CP  
; CURRENT APPLICATION NUMBER: US/09/853,526  
; CURRENT FILING DATE: 2001-05-11  
; PRIOR APPLICATION NUMBER: 09/338,907  
; PRIOR FILING DATE: 1999-06-23  
; PRIOR APPLICATION NUMBER: 08/996,306  
; PRIOR FILING DATE: 1997-12-22  
; PRIOR APPLICATION NUMBER: 60/099,658  
; PRIOR FILING DATE: 1998-09-09  
; PRIOR APPLICATION NUMBER: 09/218,207  
; PRIOR FILING DATE: 1998-12-22  
; NUMBER OF SEQ ID NOS: 578  
; SOFTWARE: Patent.pm  
; SEQ ID NO 5  
; LENGTH: 364  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-853-526-5

Query Match 100.0%; Score 364; DB 22; Length 364;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLISLVLHHTYSMRYLLPSVVLGTAPTYVLAWGVWRLLSAFUPARYOALDRLXYCVOS 60  
Db 1 MLISLVLHHTYSMRYLLPSVVLGTAPTYVLAWGVWRLLSAFUPARYOALDRLXYCVOS 60  
QY 61 MVLFFENNTGVOILLXGDLPPKKNENIYLANHSTVDMIVADILAIRONALGHVRYVLK 120  
Db 61 MVLFFENNTGVOILLXGDLPPKKNENIYLANHSTVDMIVADILAIRONALGHVRYVLK 120  
QY 121 EGLKWLPLYGCVFAOHGIIYVVKRSKAFNEKEMRNKLQSYVDAGTPMYLVIFPEGTRYNPE 180  
Db 121 EGLKWLPLYGCVFAOHGIIYVVKRSKAFNEKEMRNKLQSYVDAGTPMYLVIFPEGTRYNPE 180  
QY 181 QTKVLSASQAFAAORGLAVLKHVLTFRKATNHAFCMKNYLDAIYDVTVYVEGKDDGQ 240  
Db 181 QTKVLSASQAFAAORGLAVLKHVLTFRKATNHAFCMKNYLDAIYDVTVYVEGKDDGQ 240  
QY 241 RRESPTMTFELCKECPKIHIDRIDKDVPEOEHRMRHLHERFEIKDKMLIEFYESPD 300  
Db 241 RRESPTMTFELCKECPKIHIDRIDKDVPEOEHRMRHLHERFEIKDKMLIEFYESPD 300  
QY 301 PERRRRPCKSVNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTWYTWIYGTLLGCLMW 360  
Db 301 PERRRRPCKSVNSKLSIKKTLPSMLISGLTAGMLMTDAGRKLVTWYTWIYGTLLGCLMW 360  
QY 361 TTKA 364  
Db 361 TTKA 364

RESULT 4  
US-09-901-484A-5  
; Sequence 5, Application US/09901484A

```

; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate Cancer Gene
; FILE REFERENCE: GEN-T11XC3D2
; CURRENT APPLICATION NUMBER: US/09/901,484A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 08/996,306
; PRIOR FILING DATE: 1997-12-22
; PRIOR APPLICATION NUMBER: US 60/099,658
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: US 09/218,207
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: US 09/338,907
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: US 09/853,526
; PRIOR FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-901-484A-5

```

```

Query Match      100.0%; Score 364; DB 23; Length 364;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MLTSLVHTYSMRYLLPSVVLGTAPTYVLAAGVWRLLSAFLPARFYQALDDRLCYGQS 60
DB 1 MLTSLVHTYSMRYLLPSVVLGTAPTYVLAAGVWRLLSAFLPARFYQALDDRLCYGQS 60
QY 61 MVLEFFENTYGVQILLYGDLPRKNENIYLANHOSYVDWIVADILAIRONALGHVRYL 120
DB 61 MVLEFFENTYGVQILLYGDLPRKNENIYLANHOSYVDWIVADILAIRONALGHVRYL 120
QY 121 EGIKWLPLYGCTYPAOHGCIYVKSAPFNEKEMRNKLOSVDAGTPLYLVIFPEGSTRYNPE 180
DB 121 EGIKWLPLYGCTYPAOHGCIYVKSAPFNEKEMRNKLOSVDAGTPLYLVIFPEGSTRYNPE 180
QY 181 QTVLASAQAFAORGLAVLKHVLPRIKATHVAFDCMKNYLDAIDYDVTVVYEGKDDGGQ 240
DB 181 QTVLASAQAFAORGLAVLKHVLPRIKATHVAFDCMKNYLDAIDYDVTVVYEGKDDGGQ 240
QY 241 RRESPTMTEFLCKECPRIHIDRIDKDVPEOEHMRRLHREFEIKDKMLIEFYESPD 300
DB 241 RRESPTMTEFLCKECPRIHIDRIDKDVPEOEHMRRLHREFEIKDKMLIEFYESPD 300
QY 301 PERRKRPFGKSVNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLTVNTWITYGTLGCLMW 360
DB 301 PERRKRPFGKSVNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLTVNTWITYGTLGCLMW 360
QY 361 TIRA 364
DB 361 TIRA 364

```

```

RESULT 5
US-60-099-658-5
; Sequence 5, Application US/60099658
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Chumakov, Ilya
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; NUMBER OF SEQUENCES: 99
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbé, Martens, Olsson & Bear

```

```

; STREET: 501 West Broadway
; City: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/099,658
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelson, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: GENSET.018APR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 364 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: potential protein
; LOCATION: 1..364
US-60-099-658-5

```

```

Query Match      100.0%; Score 364; DB 26; Length 364;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 364; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MLTSLVHTYSMRYLLPSVVLGTAPTYVLAAGVWRLLSAFLPARFYQALDDRLCYGQS 60
DB 1 MLTSLVHTYSMRYLLPSVVLGTAPTYVLAAGVWRLLSAFLPARFYQALDDRLCYGQS 60
QY 61 MVLEFFENTYGVQILLYGDLPRKNENIYLANHOSYVDWIVADILAIRONALGHVRYL 120
DB 61 MVLEFFENTYGVQILLYGDLPRKNENIYLANHOSYVDWIVADILAIRONALGHVRYL 120
QY 121 EGIKWLPLYGCTYPAOHGCIYVKSAPFNEKEMRNKLOSVDAGTPLYLVIFPEGSTRYNPE 180
DB 121 EGIKWLPLYGCTYPAOHGCIYVKSAPFNEKEMRNKLOSVDAGTPLYLVIFPEGSTRYNPE 180
QY 181 QTVLASAQAFAORGLAVLKHVLPRIKATHVAFDCMKNYLDAIDYDVTVVYEGKDDGGQ 240
DB 181 QTVLASAQAFAORGLAVLKHVLPRIKATHVAFDCMKNYLDAIDYDVTVVYEGKDDGGQ 240
QY 241 RRESPTMTEFLCKECPRIHIDRIDKDVPEOEHMRRLHREFEIKDKMLIEFYESPD 300
DB 241 RRESPTMTEFLCKECPRIHIDRIDKDVPEOEHMRRLHREFEIKDKMLIEFYESPD 300
QY 301 PERRKRPFGKSVNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLTVNTWITYGTLGCLMW 360
DB 301 PERRKRPFGKSVNSKLSIKKTLPMSMLISGLTAGMLMTDAGRKLTVNTWITYGTLGCLMW 360
QY 361 TIRA 364
DB 361 TIRA 364

```

```

RESULT 6
PCT-US01-01431-59
; Sequence 59, Application PC/TUS0101431
; GENERAL INFORMATION:

```

APPLICANT: Human Genome Sciences, Inc.  
TITLE OF INVENTION: 17 human secreted proteins  
FILE REFERENCE: PS723PCT  
CURRENT APPLICATION NUMBER: PCT/US01/01431  
CURRENT FILING DATE: 2001-01-17  
PRIOR APPLICATION NUMBER: 60/179,065  
PRIOR FILING DATE: 2000-01-31  
PRIOR APPLICATION NUMBER: 60/180,628  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: 60/231,968  
PRIOR FILING DATE: 2000-09-12  
NUMBER OF SEQ ID NOS: 97  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 59  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
PCT-US01-01431-59

Query Match 97.0%; Score 353; DB 1; Length 353;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 MRYLLPSVLLGTAPTYVLAWGVRLSAPLPARFYQALDRLCYVQSWLFFFEYNTG 71  
DB 1 MRYLLPSVLLGTAPTYVLAWGVRLSAPLPARFYQALDRLCYVQSWLFFFEYNTG 60  
QY 72 VOILLYDGLPKKNENIYYLANHSTVDMIVADILAIQNALGHRVYKLGKMLPLYGC 131  
DB 61 VOILLYDGLPKKNENIYYLANHSTVDMIVADILAIQNALGHRVYKLGKMLPLYGC 120  
QY 132 YFAOHGIIYVRSKAKFNEKEMRNKLSQSYVDAGTPMYLVITPEGTRVNPEDTKVLSAQAF 191  
DB 121 YFAOHGIIYVRSKAKFNEKEMRNKLSQSYVDAGTPMYLVITPEGTRVNPEDTKVLSAQAF 180  
QY 192 AAQGLAVLKHVLTPIKATHVAFDCMKNYLDIYDVTVYVEGKDDGQRRSEPTMTTEFL 251  
DB 181 AAQGLAVLKHVLTPIKATHVAFDCMKNYLDIYDVTVYVEGKDDGQRRSEPTMTTEFL 240  
QY 252 CKECPKIHIDRIKDKDVPEQEHMRWMLHERFEIKDKMLIEFYESPDERRRRRPFPGKS 311  
DB 241 CKECPKIHIDRIKDKDVPEQEHMRWMLHERFEIKDKMLIEFYESPDERRRRRPFPGKS 300  
QY 312 VNSKLSIKTKLPMSMLISGLTAGMLMTDAGRKLYVNTWYIGTLGCLMTVITKA 364  
DB 301 VNSKLSIKTKLPMSMLISGLTAGMLMTDAGRKLYVNTWYIGTLGCLMTVITKA 353

RESULT 7  
PCT-US01-01431-79  
Sequence 79, Application PC/TUS0101431  
GENERAL INFORMATION:  
APPLICANT: Human Genome Sciences, Inc.  
TITLE OF INVENTION: 17 human secreted proteins  
FILE REFERENCE: PS723PCT  
CURRENT APPLICATION NUMBER: PCT/US01/01431  
CURRENT FILING DATE: 2001-01-17  
PRIOR APPLICATION NUMBER: 60/179,065  
PRIOR FILING DATE: 2000-01-31  
PRIOR APPLICATION NUMBER: 60/180,628  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: 60/231,968  
PRIOR FILING DATE: 2000-09-12  
NUMBER OF SEQ ID NOS: 97  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 79  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
PCT-US01-01431-79

Query Match 97.0%; Score 353; DB 1; Length 353;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 MRYLLPSVLLGTAPTYVLAWGVRLSAPLPARFYQALDRLCYVQSWLFFFEYNTG 71  
DB 1 MRYLLPSVLLGTAPTYVLAWGVRLSAPLPARFYQALDRLCYVQSWLFFFEYNTG 60  
QY 72 VOILLYDGLPKKNENIYYLANHSTVDMIVADILAIQNALGHRVYKLGKMLPLYGC 131  
DB 61 VOILLYDGLPKKNENIYYLANHSTVDMIVADILAIQNALGHRVYKLGKMLPLYGC 120  
QY 132 YFAOHGIIYVRSKAKFNEKEMRNKLSQSYVDAGTPMYLVITPEGTRVNPEDTKVLSAQAF 191  
DB 121 YFAOHGIIYVRSKAKFNEKEMRNKLSQSYVDAGTPMYLVITPEGTRVNPEDTKVLSAQAF 180  
QY 192 AAQGLAVLKHVLTPIKATHVAFDCMKNYLDIYDVTVYVEGKDDGQRRSEPTMTTEFL 251  
DB 181 AAQGLAVLKHVLTPIKATHVAFDCMKNYLDIYDVTVYVEGKDDGQRRSEPTMTTEFL 240  
QY 252 CKECPKIHIDRIKDKDVPEQEHMRWMLHERFEIKDKMLIEFYESPDERRRRRPFPGKS 311  
DB 241 CKECPKIHIDRIKDKDVPEQEHMRWMLHERFEIKDKMLIEFYESPDERRRRRPFPGKS 300  
QY 312 VNSKLSIKTKLPMSMLISGLTAGMLMTDAGRKLYVNTWYIGTLGCLMTVITKA 364  
DB 301 VNSKLSIKTKLPMSMLISGLTAGMLMTDAGRKLYVNTWYIGTLGCLMTVITKA 353

RESULT 8  
PCT-US01-11988-1474  
Sequence 1474, Application PC/TUS0111988  
GENERAL INFORMATION:  
APPLICANT: Human Genome Sciences, Inc.  
TITLE OF INVENTION: Albumin Fusion Proteins  
FILE REFERENCE: PF546PCT  
CURRENT APPLICATION NUMBER: PCT/US01/11988  
CURRENT FILING DATE: 2001-01-12  
PRIOR APPLICATION NUMBER: 60/229,358  
PRIOR FILING DATE: 2000-04-12  
PRIOR APPLICATION NUMBER: 60/256,931  
PRIOR FILING DATE: 2000-12-21  
PRIOR APPLICATION NUMBER: 60/199,384  
PRIOR FILING DATE: 2000-04-25  
NUMBER OF SEQ ID NOS: 2267  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1474  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Homo sapiens  
PCT-US01-11988-1474

Query Match 97.0%; Score 353; DB 1; Length 353;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 MRYLLPSVLLGTAPTYVLAWGVRLSAPLPARFYQALDRLCYVQSWLFFFEYNTG 71  
DB 1 MRYLLPSVLLGTAPTYVLAWGVRLSAPLPARFYQALDRLCYVQSWLFFFEYNTG 60  
QY 72 VOILLYDGLPKKNENIYYLANHSTVDMIVADILAIQNALGHRVYKLGKMLPLYGC 131  
DB 61 VOILLYDGLPKKNENIYYLANHSTVDMIVADILAIQNALGHRVYKLGKMLPLYGC 120  
QY 132 YFAOHGIIYVRSKAKFNEKEMRNKLSQSYVDAGTPMYLVITPEGTRVNPEDTKVLSAQAF 191  
DB 121 YFAOHGIIYVRSKAKFNEKEMRNKLSQSYVDAGTPMYLVITPEGTRVNPEDTKVLSAQAF 180  
QY 192 AAQGLAVLKHVLTPIKATHVAFDCMKNYLDIYDVTVYVEGKDDGQRRSEPTMTTEFL 251  
DB 181 AAQGLAVLKHVLTPIKATHVAFDCMKNYLDIYDVTVYVEGKDDGQRRSEPTMTTEFL 240

```
Qy 252 CKECPKIHIDRIDKKDVEEEOEHMRRLHEFEIKDKMLIEFYSPDERRRKRPFGKS 311
      |||||||
Db 241 CKECPKIHIDRIDKKDVEEEOEHMRRLHEFEIKDKMLIEFYSPDERRRKRPFGKS 300
Qy 312 VNSKLSIKKTLPSSMLISGLTAGMLMTDAGRKLYVNTWYIYGTLLGCLMWYTIKA 364
      |||||||
Db 301 VNSKLSIKKTLPSSMLISGLTAGMLMTDAGRKLYVNTWYIYGTLLGCLMWYTIKA 353
```

```
RESULT 9
PCT-US01-11988-1475
; Sequence 1475, Application PC/TUS0111988
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: PCT/US01/11988
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 1475
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-11988-1475
```

```
Query Match          97.0%: Score 353; DB 1; Length 353;
Best Local Similarity 100.0%: Pred. No. 0;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 12 MRYLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQSQWVLEFFENYTG 71
      |||||||
Db 1 MRYLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQSQWVLEFFENYTG 60
Qy 72 VOILLYGDLPKKNENIITYLANHSTVDWIADILAIQNALGHVRYLKEGLKWLPLYGC 131
      |||||||
Db 61 VOILLYGDLPKKNENIITYLANHSTVDWIADILAIQNALGHVRYLKEGLKWLPLYGC 120
Qy 132 YFAOHGIIYVKRSKAFNEKEMRNKLOSVDAGPMTLVIFPESTRYNPEOTKYLSSQAF 191
      |||||||
Db 121 YFAOHGIIYVKRSKAFNEKEMRNKLOSVDAGPMTLVIFPESTRYNPEOTKYLSSQAF 180
Qy 192 AAORGLAVLKHVLTPIKATHVAFDCMKNYLDAIYDVTVYVEGKDDGGORRESPTMTTEFL 251
      |||||||
Db 181 AAORGLAVLKHVLTPIKATHVAFDCMKNYLDAIYDVTVYVEGKDDGGORRESPTMTTEFL 240
Qy 252 CKECPKIHIDRIDKKDVEEEOEHMRRLHEFEIKDKMLIEFYSPDERRRKRPFGKS 311
      |||||||
Db 241 CKECPKIHIDRIDKKDVEEEOEHMRRLHEFEIKDKMLIEFYSPDERRRKRPFGKS 300
Qy 312 VNSKLSIKKTLPSSMLISGLTAGMLMTDAGRKLYVNTWYIYGTLLGCLMWYTIKA 364
      |||||||
Db 301 VNSKLSIKKTLPSSMLISGLTAGMLMTDAGRKLYVNTWYIYGTLLGCLMWYTIKA 353
```

```
RESULT 10
US-09-833-245-1474
; Sequence 1474, Application US/09833245
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
```

```
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 1474
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-245-1474
```

```
Query Match          97.0%: Score 353; DB 22; Length 353;
Best Local Similarity 100.0%: Pred. No. 0;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 12 MRYLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQSQWVLEFFENYTG 71
      |||||||
Db 1 MRYLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQSQWVLEFFENYTG 60
Qy 72 VOILLYGDLPKKNENIITYLANHSTVDWIADILAIQNALGHVRYLKEGLKWLPLYGC 131
      |||||||
Db 61 VOILLYGDLPKKNENIITYLANHSTVDWIADILAIQNALGHVRYLKEGLKWLPLYGC 120
Qy 132 YFAOHGIIYVKRSKAFNEKEMRNKLOSVDAGPMTLVIFPESTRYNPEOTKYLSSQAF 191
      |||||||
Db 121 YFAOHGIIYVKRSKAFNEKEMRNKLOSVDAGPMTLVIFPESTRYNPEOTKYLSSQAF 180
Qy 192 AAORGLAVLKHVLTPIKATHVAFDCMKNYLDAIYDVTVYVEGKDDGGORRESPTMTTEFL 251
      |||||||
Db 181 AAORGLAVLKHVLTPIKATHVAFDCMKNYLDAIYDVTVYVEGKDDGGORRESPTMTTEFL 240
Qy 252 CKECPKIHIDRIDKKDVEEEOEHMRRLHEFEIKDKMLIEFYSPDERRRKRPFGKS 311
      |||||||
Db 241 CKECPKIHIDRIDKKDVEEEOEHMRRLHEFEIKDKMLIEFYSPDERRRKRPFGKS 300
Qy 312 VNSKLSIKKTLPSSMLISGLTAGMLMTDAGRKLYVNTWYIYGTLLGCLMWYTIKA 364
      |||||||
Db 301 VNSKLSIKKTLPSSMLISGLTAGMLMTDAGRKLYVNTWYIYGTLLGCLMWYTIKA 353
```

```
RESULT 11
US-09-833-245-1475
; Sequence 1475, Application US/09833245
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 1475
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-245-1475
```

```
Query Match          97.0%: Score 353; DB 22; Length 353;
Best Local Similarity 100.0%: Pred. No. 0;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 MRYLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQSQWVLEFFENYTG 71
      |||||||
Db 1 MRYLPSVVLGTAPTYVLAMGWRLLSAFLPARFYQALDDRLCYQSQWVLEFFENYTG 60
```

QY 72 VOILLYDLPKKNENIYLANHSTVDWIVADILAIROMALGHVRYVLEKESLKWLPYGC 131  
DB 61 VOILLYDLPKKNENIYLANHSTVDWIVADILAIROMALGHVRYVLEKESLKWLPYGC 120  
QY 132 YFNOHGIIYVRSKAKFEKEMKRNKLOSIVDAGTPMYIVIEPEGRIYRPEOTKVSASQAF 191  
DB 121 YFNOHGIIYVRSKAKFEKEMKRNKLOSIVDAGTPMYIVIEPEGRIYRPEOTKVSASQAF 180  
QY 192 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDATYDVTVYEGKDDGSGORRESPTMEFL 251  
DB 181 AAOGLAVLKHVLTLPRIKATHVAFDCMKNYLDATYDVTVYEGKDDGSGORRESPTMEFL 240  
QY 252 CKECPRKIHIDRIKDKVPEDEHMRMWLHEREFIDKMLIEFYESPDERRKRFPKGS 311  
DB 241 CKECPRKIHIDRIKDKVPEDEHMRMWLHEREFIDKMLIEFYESPDERRKRFPKGS 300  
QY 312 VNSKLSIKTKLPMSLLISGLTAGMLMDAGRKLYVNTWYIGTLGCLMWTIKA 364  
DB 301 VNSKLSIKTKLPMSLLISGLTAGMLMDAGRKLYVNTWYIGTLGCLMWTIKA 353

## RESULT 12

US-09-853-526-4

Sequence 4, Application US/09853526

GENERAL INFORMATION:

APPLICANT: Cohen, Daniel

APPLICANT: Blumenfeld, Marta

APPLICANT: Ilyu, Chumakov

APPLICANT: Bougueleret, Lydie

TITLE OF INVENTION: PROSTATE CANCER GENE

FILE REFERENCE: GENSET.18CP1CP

CURRENT APPLICATION NUMBER: US/09/853,526

PRIORITY FILING DATE: 2001-05-11

PRIORITY APPLICATION NUMBER: 09/338,907

PRIORITY FILING DATE: 1999-06-23

PRIORITY APPLICATION NUMBER: 08/996,306

PRIORITY FILING DATE: 1997-12-22

PRIORITY APPLICATION NUMBER: 60/099,658

PRIORITY FILING DATE: 1998-09-09

PRIORITY APPLICATION NUMBER: 09/218,207

PRIORITY FILING DATE: 1998-12-22

NUMBER OF SEQ ID NOS: 578

SOFTWARE: Patent.pm

SEQ ID NO 4

LENGTH: 353

TYPE: PR

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: HELIX

LOCATION: 1..33

OTHER INFORMATION: Rao and Argos identification method, potential helix

NAME/KEY: HELIX

LOCATION: 4..20

OTHER INFORMATION: Klein, Kanehisa and Delisi identification method, potential

NAME/KEY: HELIX

LOCATION: 4..24

OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F

NAME/KEY: MYRISTATE

LOCATION: 12..16

OTHER INFORMATION: Prosite match

NAME/KEY: HELIX

LOCATION: 50..70

OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F

NAME/KEY: CARBOHYD

LOCATION: 57..59

OTHER INFORMATION: Prosite match

NAME/KEY: HELIX

LOCATION: 76..96

OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method, F

NAME/KEY: PHOSPHORYLATION  
LOCATION: 78  
OTHER INFORMATION: potential Tyrosine kinase site, Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 84  
OTHER INFORMATION: potential caseine kinase II site, Prosite match  
NAME/KEY: SITE  
LOCATION: 94..115  
OTHER INFORMATION: potential leucine zipper site, Prosite match  
NAME/KEY: MYRISTATE  
LOCATION: 119..123  
OTHER INFORMATION: potential site, Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 133  
OTHER INFORMATION: potential protein kinase C, Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 147  
OTHER INFORMATION: potential caseine kinase II site, Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 194  
OTHER INFORMATION: potential protein kinase C, Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 215  
OTHER INFORMATION: potential Tyrosine kinase site, Prosite match  
NAME/KEY: SULFATATION  
LOCATION: 221  
OTHER INFORMATION: Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 233  
OTHER INFORMATION: potential cAMP and cGMP dependant protein kinase site,  
OTHER INFORMATION: match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 235  
OTHER INFORMATION: potential caseine kinase II site, Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 306  
OTHER INFORMATION: potential protein kinase C, Prosite match  
NAME/KEY: HELIX  
LOCATION: 310..330  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
OTHER INFORMATION: helix  
NAME/KEY: MYRISTATE  
LOCATION: 319..323  
OTHER INFORMATION: Prosite match  
NAME/KEY: MYRISTATE  
LOCATION: 323..327  
OTHER INFORMATION: Prosite match  
NAME/KEY: AMIDATION  
LOCATION: 329  
OTHER INFORMATION: Prosite match  
NAME/KEY: HELIX  
LOCATION: 333..353  
OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,  
OTHER INFORMATION: helix  
NAME/KEY: MYRISTATE  
LOCATION: 341..345  
OTHER INFORMATION: Prosite match  
NAME/KEY: PHOSPHORYLATION  
LOCATION: 350  
OTHER INFORMATION: potential protein kinase C, Prosite match

US-09-853-526-4

## Query Match

97.0%; Score 353; DB 22; Length 353;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 MYLLPSVVLGTAPTYVLAAGVWRLLSAFLPARFYQALDDRLCYVQSMVLFEEENYTG 71  
DB 1 MYLLPSVVLGTAPTYVLAAGVWRLLSAFLPARFYQALDDRLCYVQSMVLFEEENYTG 60  
QY 72 VOILLYDLPKKNENIYLANHSTVDWIVADILAIROMALGHVRYVLEKESLKWLPYGC 131  
|||||

```

Db 61 VOILLVGLDLPKKNENITILANHOSTVDWIVADILAI RONALGHVRYVLEKGLKWLPLYGC 120
? LOCATION: (78)..(78)
? OTHER INFORMATION: PHOSPHORYLATION, potential Tyrosine kinase site, Prosite match
Oy 132 YFAOHGIYVKSRAKNEKEMRNKLOSVDAGT PMVLTFPECTRYNPEOTKVLASOAF 191
? NAME/KEY: MOD_RES
? LOCATION: (84)..(84)
Db 121 YRAOHGIYVKSRAKNEKEMRNKLOSVDAGT PMVLTFPECTRYNPEOTKVLASOAF 180
? OTHER INFORMATION: PHOSPHORYLATION, potential caseline kinase II site, Prosite mat
? NAME/KEY: SITE
? LOCATION: (94)..(115)
Oy 192 AAORGLAVLKHVLPTRIKATHAFCDMKNYLDAIYDVTVVYEGKDDGGORRESPTMTEFL 251
? OTHER INFORMATION: Potential leucine zipper site, Prosite match
? NAME/KEY: LIPID
Db 181 AAORGLAVLKHVLPTRIKATHAFCDMKNYLDAIYDVTVVYEGKDDGGORRESPTMTEFL 240
? LOCATION: (119)..(123)
? OTHER INFORMATION: MYRISTATE, Prosite match
? NAME/KEY: SITE
Oy 252 CCECPRIHIDRIDKDVPEDEEHMRRLVHERFEIKDKMLIEFYSPDPERKRRPGKS 311
? NAME/KEY: MOD_RES
? LOCATION: (133)..(133)
Db 241 CCECPRIHIDRIDKDVPEDEEHMRRLVHERFEIKDKMLIEFYSPDPERKRRPGKS 300
? OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match
? NAME/KEY: MOD_RES
? LOCATION: (147)..(147)
Oy 312 VNSKISIKTLPMSLTLSGLTAGMLTFDAGRKLVTWTYGTLLGLWTYTKA 364
? OTHER INFORMATION: PHOSPHORYLATION, potential caseline kinase II, Prosite match
? NAME/KEY: MOD_RES
Db 301 VNSKISIKTLPMSLTLSGLTAGMLTFDAGRKLVTWTYGTLLGLWTYTKA 353
? LOCATION: (194)..(194)
? OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match
? NAME/KEY: MOD_RES
? LOCATION: (215)..(215)
? OTHER INFORMATION: PHOSPHORYLATION, potential Tyrosine kinase site, Prosite match
? NAME/KEY: MOD_RES
? LOCATION: (221)..(221)
? OTHER INFORMATION: SULFATATION, Prosite match
? NAME/KEY: MOD_RES
? LOCATION: (233)..(233)
? OTHER INFORMATION: PHOSPHORYLATION, potential cAMP and cGMP dependant protein kin
? OTHER INFORMATION: site, Prosite match
? NAME/KEY: MOD_RES
? LOCATION: (235)..(235)
? OTHER INFORMATION: PHOSPHORYLATION, potential caseline kinase II site, Prosite mat
? NAME/KEY: MOD_RES
? LOCATION: (306)..(306)
? OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match
? NAME/KEY: HELIX
? LOCATION: (310)..(330)
? OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
? OTHER INFORMATION: potential helix
? NAME/KEY: LIPID
? LOCATION: (319)..(323)
? OTHER INFORMATION: MYRISTATE, Prosite match
? NAME/KEY: LIPID
? LOCATION: (323)..(327)
? OTHER INFORMATION: MYRISTATE, Prosite match
? NAME/KEY: MOD_RES
? LOCATION: (329)..(329)
? OTHER INFORMATION: AMIDATION, Prosite match
? NAME/KEY: HELIX
? LOCATION: (333)..(353)
? OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
? OTHER INFORMATION: potential helix
? NAME/KEY: LIPID
? LOCATION: (341)..(345)
? OTHER INFORMATION: MYRISTATE, Prosite match
? NAME/KEY: MOD_RES
? LOCATION: (350)..(350)
? OTHER INFORMATION: PHOSPHORYLATION, potential protein kinase C, Prosite match
US-09-901-484A-4

RESULT 13
US-09-901-484A-4
? Sequence 4, Application US/09901484A
? GENERAL INFORMATION:
? APPLICANT: Cohen, Daniel
? APPLICANT: Blumenfeld, Marta
? APPLICANT: Chumakov, Ilya
? APPLICANT: Bounelelet, Lydie
? TITLE OF INVENTION: Prostate Cancer Gene
? FILE REFERENCE: GEN-T11XC3D2
? CURRENT FILING DATE: 2001-07-09
? PRIOR APPLICATION NUMBER: US/09/901,484A
? PRIOR FILING DATE: 1997-12-22
? PRIOR APPLICATION NUMBER: US 60/099,658
? PRIOR FILING DATE: 1998-09-09
? PRIOR APPLICATION NUMBER: US 09/218,207
? PRIOR FILING DATE: 1998-12-22
? PRIOR APPLICATION NUMBER: US 09/338,907
? PRIOR FILING DATE: 1999-06-23
? PRIOR APPLICATION NUMBER: US 09/853,526
? NUMBER OF SEQ ID NOS: 578
? SOFTWARE: PatentIn version 3.1
? SEQ ID NO 4
? LENGTH: 353
? TYPE: PRT
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: HELIX
? LOCATION: (1)..(33)
? OTHER INFORMATION: Rao and Argos identification method, potential helix
? NAME/KEY: HELIX
? LOCATION: (4)..(20)
? OTHER INFORMATION: Klein, Kanehisa and Delisi identification method, potential helix
? NAME/KEY: HELIX
? LOCATION: (4)..(24)
? OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
? NAME/KEY: LIPID
? LOCATION: (12)..(16)
? OTHER INFORMATION: MYRISTATE, Prosite match
? NAME/KEY: HELIX
? LOCATION: (50)..(70)
? OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
? NAME/KEY: CARBOHYD
? LOCATION: (57)..(59)
? OTHER INFORMATION: Prosite match
? NAME/KEY: HELIX
? LOCATION: (76)..(96)
? OTHER INFORMATION: Eisenberg, Schwarz, Komarony, Wall identification method,
? OTHER INFORMATION: potential helix
? NAME/KEY: MOD_RES

```

```

Query Match          97.0%; Score 353; DB 23; Length 353;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 12 MRVLPSVVLGTAPTYVLAAGVWRLLSAFLPARYQALDDRLCYCYQSMVLFPPENYTG 71
Db 1 MRVLPSVVLGTAPTYVLAAGVWRLLSAFLPARYQALDDRLCYCYQSMVLFPPENYTG 60
Oy 72 VOILLVGLDLPKKNENITILANHOSTVDWIVADILAI RONALGHVRYVLEKGLKWLPLYGC 131
Db 61 VOILLVGLDLPKKNENITILANHOSTVDWIVADILAI RONALGHVRYVLEKGLKWLPLYGC 120

```

```
QY 132 YFAOHGSIYKRSKAKFNEKEMRNKLSQSYVDAGTPMYLVIFPEGTRYNPEQTKVLSASQAF 191
|
|
|
Db 121 YFAOHGSIYKRSKAKFNEKEMRNKLSQSYVDAGTPMYLVIFPEGTRYNPEQTKVLSASQAF 180
|
|
|
QY 192 AAORGLAVLKHVLTLPRIKATHVAFDCMKNTLDIYDTVYVEGKDDGGORRESPTMTTEFL 251
|
|
|
Db 181 AAORGLAVLKHVLTLPRIKATHVAFDCMKNTLDIYDTVYVEGKDDGGORRESPTMTTEFL 240
|
|
|
QY 252 CKCECPKIHIDRIDKKDVEEEOEHMRMWLHERFEIKDKMLIEFYESPDEPRKRRPFGKS 311
|
|
|
Db 241 CKCECPKIHIDRIDKKDVEEEOEHMRMWLHERFEIKDKMLIEFYESPDEPRKRRPFGKS 300
|
|
|
QY 312 VNSKLSIKTKLPMSMLISGLTAGMLMTDAGRKLVTWMTWYIGTLGCLMTWTIKA 364
|
|
|
Db 301 VNSKLSIKTKLPMSMLISGLTAGMLMTDAGRKLVTWMTWYIGTLGCLMTWTIKA 353
```

```
RESULT 14
US-09-915-582-59
; Sequence 59, Application US/09915582
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 17 Human Secreted Proteins
; FILE REFERENCE: PS723P1
; CURRENT APPLICATION NUMBER: US/09/915,582
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: PCT/US01/01431
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/231,968
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 59
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-915-582-59
```

```
Query Match 97.0%; Score 353; DB 23; Length 353;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 12 MRYLLPSVVLGTAPTYVVLAMGWRLLSAFLPARFYQALDDRLCYQOSVYVLEFFENYTG 71
|
|
|
Db 1 MRYLLPSVVLGTAPTYVVLAMGWRLLSAFLPARFYQALDDRLCYQOSVYVLEFFENYTG 60
|
|
|
QY 72 VOILLGDLPRKKNENITYLANHOSTVDWIVADILAIROMALGHVRYVLEKGLMWLPLYGC 131
|
|
|
Db 61 VOILLGDLPRKKNENITYLANHOSTVDWIVADILAIROMALGHVRYVLEKGLMWLPLYGC 120
|
|
|
QY 132 YFAOHGSIYKRSKAKFNEKEMRNKLSQSYVDAGTPMYLVIFPEGTRYNPEQTKVLSASQAF 191
|
|
|
Db 121 YFAOHGSIYKRSKAKFNEKEMRNKLSQSYVDAGTPMYLVIFPEGTRYNPEQTKVLSASQAF 180
|
|
|
QY 192 AAORGLAVLKHVLTLPRIKATHVAFDCMKNTLDIYDTVYVEGKDDGGORRESPTMTTEFL 251
|
|
|
Db 181 AAORGLAVLKHVLTLPRIKATHVAFDCMKNTLDIYDTVYVEGKDDGGORRESPTMTTEFL 240
|
|
|
QY 252 CKCECPKIHIDRIDKKDVEEEOEHMRMWLHERFEIKDKMLIEFYESPDEPRKRRPFGKS 311
|
|
|
Db 241 CKCECPKIHIDRIDKKDVEEEOEHMRMWLHERFEIKDKMLIEFYESPDEPRKRRPFGKS 300
|
|
|
QY 312 VNSKLSIKTKLPMSMLISGLTAGMLMTDAGRKLVTWMTWYIGTLGCLMTWTIKA 364
|
|
|
Db 301 VNSKLSIKTKLPMSMLISGLTAGMLMTDAGRKLVTWMTWYIGTLGCLMTWTIKA 353
```

```
RESULT 15
US-09-915-582-79
; Sequence 79, Application US/09915582
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 17 Human Secreted Proteins
; FILE REFERENCE: PS723P1
; CURRENT APPLICATION NUMBER: US/09/915,582
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: PCT/US01/01431
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/231,968
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 79
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-915-582-79
```

```
Query Match 97.0%; Score 353; DB 23; Length 353;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 12 MRYLLPSVVLGTAPTYVVLAMGWRLLSAFLPARFYQALDDRLCYQOSVYVLEFFENYTG 71
|
|
|
Db 1 MRYLLPSVVLGTAPTYVVLAMGWRLLSAFLPARFYQALDDRLCYQOSVYVLEFFENYTG 60
|
|
|
QY 72 VOILLGDLPRKKNENITYLANHOSTVDWIVADILAIROMALGHVRYVLEKGLMWLPLYGC 131
|
|
|
Db 61 VOILLGDLPRKKNENITYLANHOSTVDWIVADILAIROMALGHVRYVLEKGLMWLPLYGC 120
|
|
|
QY 132 YFAOHGSIYKRSKAKFNEKEMRNKLSQSYVDAGTPMYLVIFPEGTRYNPEQTKVLSASQAF 191
|
|
|
Db 121 YFAOHGSIYKRSKAKFNEKEMRNKLSQSYVDAGTPMYLVIFPEGTRYNPEQTKVLSASQAF 180
|
|
|
QY 192 AAORGLAVLKHVLTLPRIKATHVAFDCMKNTLDIYDTVYVEGKDDGGORRESPTMTTEFL 251
|
|
|
Db 181 AAORGLAVLKHVLTLPRIKATHVAFDCMKNTLDIYDTVYVEGKDDGGORRESPTMTTEFL 240
|
|
|
QY 252 CKCECPKIHIDRIDKKDVEEEOEHMRMWLHERFEIKDKMLIEFYESPDEPRKRRPFGKS 311
|
|
|
Db 241 CKCECPKIHIDRIDKKDVEEEOEHMRMWLHERFEIKDKMLIEFYESPDEPRKRRPFGKS 300
|
|
|
QY 312 VNSKLSIKTKLPMSMLISGLTAGMLMTDAGRKLVTWMTWYIGTLGCLMTWTIKA 364
|
|
|
Db 301 VNSKLSIKTKLPMSMLISGLTAGMLMTDAGRKLVTWMTWYIGTLGCLMTWTIKA 353
```

Search completed: August 28, 2002, 11:25:19  
Job time: 497 sec